医学遗传学

主　编　蔡绍京　梁素华　税青林

副主编　(按姓氏笔画为序)

马志敏　贡昌春　李学英

金　洁　施旭东

编　者　(按姓氏笔画为序)

马志敏(大理学院)　王秀英(徐州医学院)

刘福民(徐州医学院)　孙贝加(徐州医学院)

贡昌春(扬州大学)　李学英(遵义医学院)

余　红(泸州医学院)　张　励(大理学院)

金　洁(江苏大学)　施旭东(徐州医学院)

钱　刚(遵义医学院)　梁素华(川北医学院)

税青林(泸州医学院)　蔡绍京(徐州医学院)

霍正浩(宁夏医学院)

主　审　霍正浩　王秀英

第二军医大学出版社

前 言

为了适应高等医学教育面向21世纪课程体系改革的需要，充分体现素质教育和个性教育的思想，切实提高教学水平和教学质量，适应培养高素质、宽口径医学人才的需要，我们5省区8院校的同行们共同编写了这本《医学遗传学》。

医学遗传学是医学科学领域十分活跃的前沿学科，分子生物学方法的引入及人类基因组计划的推动，促进了医学遗传学的飞速发展，目前，医学遗传学已发展成为涉及数千种遗传性疾病的基础理论和临床实践的科学。人们对遗传病的认识已达到了新的高度，不仅对单基因遗传病和多基因遗传病的诊断、发病机制、治疗和预防达到了分子水平，而且由于显微切割、荧光原位杂交(FISH)等方法的应用，染色体病的诊断也已深入到基因片段的水平；癌基因及肿瘤抑制基因的深入研究使人们对肿瘤发生、发展的机制有了更深入的理解；遗传病、肿瘤、心血管疾病等基因治疗的各种策略为战胜这些疾病展示了光明的前景；人类基因组计划的研究进度一再提前，人类基因组DNA全序列这部天书完全解读已指日可待。这些研究成果必将引导21世纪的生物医学结出丰硕成果，造福于人类。作为培养21世纪医学人才的《医学遗传学》教材只有更新内容，才能跟上医学科学发展的步伐，才能使培养出的医学人才有较新的知识储备。我们这本《医学遗传学》正是在这样背景下产生的。

全书共16章3大部分。第1部分从细胞和分子水平介绍遗传的物质基础，这部分内容是医学遗传学的基础知识，起着承前启后的作用；第2部分是本书的主体，包括各类人类遗传病，遗传病的诊断、治疗与预防，以及肿瘤遗传学、免疫遗传学、药物遗传学、群体遗传学等；第3部分包括人类基因的研究技术和人类基因组计划两章，可供学时充足的院校选用；这两章有助于学生拓宽视野、扩大知识面，并加深对遗传病的诊断与治疗等相关内容的理解。

本书在编写时，遵循突出基本概念、基本知识和基本理论，反映医学遗传学最新研究进展的原则，力图做到既便于教、又便于学。各章之后的小结旨在

帮助学生掌握各章重点，书末附有英汉名词索引，方便查阅。

参与本书编写的各位作者，在教学、科研工作十分繁忙的情况下，牺牲了大量休息时间，对所负责章节的编写工作，精益求精、一丝不苟、数易其稿，体现了对学生的高度负责精神。本书的编写得到了徐州医学院郑葵阳副院长、基础医学部邢淑华主任及第二军医大学出版社、徐州医学院学报编辑部的大力支持；遵义医学院的领导和生物学教研室的全体教师为审稿会提供了方便，做出了积极的贡献，特此致谢。

虽然各位编者做了最大的努力，但由于水平所限，本书难免存在缺点及错误，真诚期望同行专家、使用本教材的师生及其他读者指出本书的不足之处，以便再版时修正。

蔡绍京　霍正浩

2004年8月

目 录

第一章 医学遗传学概论

第一节 医学遗传学的研究范围

医学遗传学(medical genetics)是医学与遗传学相结合、并互相渗透的一门边缘学科,是遗传学知识在医学领域的应用。它研究人类遗传性疾病的发生机制、传递规律、诊断方法以及治疗与预防措施。医学遗传学和人类遗传学都以人为研究对象。人类遗传学主要从人种和人类发展史的角度来研究人的遗传性状,如人体的形态、人种的特征;同时广泛研究形态结构、生理功能的变异,如毛发的颜色、耳的形状等,这些变异并不干扰或破坏正常的生命活动过程,故临床意义不大。医学遗传学则是从医学角度来研究人类疾病与遗传的关系,为防治遗传病及与遗传有关的疾病提供科学依据和手段,为改善人类的健康素质做出贡献。

医学遗传学是以人类遗传学为基础,借助于现代生物学的研究方法,在遗传学理论指导和实验方法被广泛采用的基础上发展起来的,至今已发展成为由众多分支学科组成的、涉及多门基础学科与临床学科的综合性学科。医学遗传学的主要分支学科包括:

1.细胞遗传学(cytogenetics) 细胞遗传学是在细胞水平的染色体遗传学说的基础上发展起来的,它从染色体的结构和行为方面研究染色体的遗传机制及其规律,研究人类染色体的结构、畸变类型、畸变发生频率及与疾病的关系。随着新技术的不断应用,细胞遗传学将对染色体的分子结构及其缺陷有更细微的认识。

2.分子遗传学(molecular genetics) 分子遗传学是用现代分子生物学技术,从基因的结构、突变、表达、调控等方面研究遗传病患者遗传物质分子水平的改变,为遗传病的基因诊断、基因治疗提供策略和手段。

3.生化遗传学(biochemical genetics) 生化遗传学应用生物化学的方法,研究遗传物质的理化性质、蛋白质的生物合成、机体代谢的调控、基因突变及其机制等。这使人们认识到分子水平的遗传物质的异常将导致分子病或先天性代谢病。近年来,随着分子生物学的理论和技术的广泛应用,人们对遗传病的本质有了更深刻的认识。

4.群体遗传学(population genetics) 群体遗传学研究群体的基因行为,探讨人类正常性状和病理性状在群体中的分布及变迁规律,辨析遗传因素和环境因素在疾病中的作用,并为遗传病的群体监控和预防制订对策和措施。

5.免疫遗传学(immunogenetics) 免疫遗传学研究免疫反应的遗传基础与遗传控制,从分子水平阐明人类免疫现象的遗传和变异规律以及与遗传有关的免疫性疾病的遗传背景,为临床实践中的输血、器官移植等提供理论基础。

6.体细胞遗传学(somatic cell genetics) 体细胞遗传学通过离体培养的体细胞研究DNA复制、基因突变、基因调控、细胞分化、个体发育、肿瘤细胞的形成机制等。体细胞遗

传学的重要研究技术是细胞融合技术，由此产生的杂交细胞在单克隆抗体的制备和基因定位上有重要作用。

7.肿瘤遗传学(cancer genetics) 肿瘤遗传学着重研究肿瘤发生、发展的遗传因素，肿瘤细胞形成、发展和转移的遗传规律。近年来，对癌基因和抗癌基因的广泛而深入的研究，开辟了肿瘤的发生机制、肿瘤的诊断与肿瘤治疗研究的新领域。

8.药物遗传学(pharmacogenetics) 药物遗传学研究药物反应个体差异的遗传学基础，在理论上阐明药物反应的遗传易感性的物质基础，在实践上为指导医生遵循用药的个体化原则提供理论根据。

总之，医学遗传学的研究范围非常广泛，而且也与医学实践密切相关。近年来，蓬勃兴起的人类基因组的研究，进一步促进了医学遗传学的发展。人们有理由相信，征服癌症、根治遗传病已不再是遥远的梦想了。

第二节　医学遗传学的发展

医学遗传学的发展历程大致经历了缓慢、快速、飞跃3个阶段。

早在古希腊希波克拉底(Hippocrates，公元前460～375)时代之前，人们认识到某些疾病可能在家庭中传递。大约1500年前，犹太教法典就有对“易出血者”的某些男性亲属免除割礼的规定，说明人们当时已认识了血友病的遗传规律，这表明当时人们对医学遗传学已有了最初步的认识。

孟德尔(Mendel，1822～1884)，作为现代遗传学的奠基者，以著名的豌豆杂交试验所得出的分离律和自由组合律，即孟德尔第一定律和孟德尔第二定律而闻名于世。其后，1910年，美国哥伦比亚大学的摩尔根(Morgan，1866～1945)和他的学生们通过果蝇的杂交试验得出了连锁律和交换律。在孟德尔、摩尔根经典遗传学理论的指引下，人们对遗传病的来源及传递方式进行了朴实的描述。

20世纪初，Garrod和Bateson首次运用孟德尔遗传定律对尿黑酸尿症的家系进行了观察研究，阐明了该病隐性遗传的遗传方式，可以认为这是医学遗传学的起始。随着遗传学的发展，医学遗传学也有了相应的进展。这个时期，人们主要在群体水平对遗传病进行调查并对不同的遗传病进行分类、描述及总结其规律。总的来说，发展的速度不快。

自20世纪50年代起，医学遗传学进入了快速发展的时期。分子病和先天性代谢病是由于结构蛋白和酶异常所引起，生化实验技术和分析方法的发展提高了分子病和遗传性代谢病的研究和临床诊断水平，所发现的病种由20世纪初的少数几种(如，尿黑酸尿症、白化病等)扩大到数百种，从生化水平上揭示了血红蛋白病、葡萄糖-6-磷酸脱氢酶(G6pD)缺乏症、苯丙酮尿症、尿黑酸尿症和高胆固醇血症等疾病的发病机制。在实际应用上，开辟了治疗某些遗传病的有效途径，苯丙酮尿症的治疗标志着这方面的重大进展。1953年，Bickel等提出，控制新生儿的苯丙氨酸摄入量可有效防止苯丙酮尿症的发展，这种饮食疗法在临床上取得了良好的治疗效果。

遗传物质的改变是遗传性疾病的本质，约有10%的遗传性疾病表现为染色体畸变。细胞遗传学的研究阐明了染色体畸变机制，发现了一大批染色体畸变综合征，对染色体缺

失、倒位、重复等形成机制有了明确的认识，特别是随着染色体显带技术、姐妹染色单体互换技术和高分辨显带技术的应用和发展，使人们能够越来越精细地识别染色体及其区带，为研究各种遗传病及遗传相关疾病的染色体畸变提供了有力工具，从染色体水平揭示了遗传性疾病及相关疾病的发病机制。

生化分析和细胞遗传学方法是医学遗传学的两种主要研究技术，这两种技术在医学中的应用加速了医学遗传学的发展，使得医学遗传学在各个领域取得了重大进展，并形成和建立了一些分支学科。如：研究电离辐射对遗传物质影响及其规律的辐射遗传学；研究基因对发育过程的控制与调节；研究基因在发育不同阶段的表达及调控机制的发育遗传学；研究基因对人类行为影响的行为遗传学；用遗传学方法研究环境因素对遗传物质的损坏及其毒理机制的毒理遗传学等。

20世纪70年代，随着限制性核酸内切酶的发现，重组DNA技术和聚合酶链反应(PCR)技术的发展和广泛运用，传统的医学遗传学发展到现代的医学分子遗传学水平，开始了飞速发展的新阶段。人们在对人类疾病的研究中，越来越清楚地认识到，只有进行基因水平的研究，才能真正找到致病的根本原因，也才能对遗传病进行最有效的防治。重组DNA技术的核心是在体外进行特定DNA片段的重组，然后在合适的宿主细胞中扩增；PCR是DNA聚合酶链式反应的简称，运用PCR技术可以在体外将DNA扩增100万倍。这两项技术可用于人类基因结构、功能的研究和疾病相关基因的克隆，揭示疾病的分子病理机制；可用于人基因组DNA文库或cDNA文库构建，制备基因探针；可用于基因表达研究，生产出临床治疗所需要的多肽药物；利用重组DNA和PCR技术，还可以将基因作为“药物”，在基因水平对遗传病、肿瘤及传染病等进行基因治疗研究。重组DNA和PCR等分子生物学技术正在对医学遗传学的各个领域产生重要影响，导致了医学遗传学领域里的一次革命，取得了突飞猛进的发展。

现将20世纪，特别是50年代以来，医学遗传学发展的大事概要总结如表1-1。

表1-1 医学遗传学发展大事概要

年份	大事概要	主要研究者	重要意义
1900	孟德尔豌豆杂交试验结果被后人总结为孟德尔定律	Mendel	奠定现代遗传学基础
1902	首次运用孟德尔定律解释尿黑酸尿症的遗传方式	Barrod, Bateson	医学遗传学起始的标志
1903	提出遗传因子(1909年改称基因)存在于染色体上	Sutton, Boveri, Johannsen	建立染色体遗传学说
1908	总结出Hardy-Weinberg定律	Hardy, Weinberg	奠定群体遗传学基础
1909	提出“多因子遗传”假说	Nilsson, Ehle	
1910	阐明连锁律和交换律	Morgan	细胞遗传学起始的标志
1926	发表《基因论》		
1941	提出“一个基因一个酶”学说	Beadle, Tatum	
1944	证明DNA是遗传物质	Avery, MacLeod, McCarty	奠定分子遗传学基础

待续

续表 1－1

年份	大事概要	主要研究者	重要意义
1949	发现镰形细胞贫血症患者的异常血红蛋白 Hb S	Pauling	提出分子病的概念
1952	证实糖原贮积病Ⅰ型是葡萄糖－6－磷酸酶缺陷所致	Cori C F, Cori G T	发现先天性代谢病
1953	发现 DNA 双螺旋结构	Watson, Crick	分子遗传学起始的标志
1956	确定人染色体数目为 46	Tjio, Levan	人类细胞遗传学开始的标志
1959	发现 Down 综合征、Turner 综合征和 Klinefelter 综合征	Lejeune, Ford, Jacobs	提出“染色体病”的概念
	发现琥珀酰胆碱高敏感个体	Vogel	提出药物遗传学概念
1960	发现 Ph 染色体	Nowell	肿瘤细胞遗传学的里程碑
1961	发现血清蛋白及红细胞中酶遗传多态性	Smithies	
1966	《人类孟德尔遗传：人类基因和遗传病目录》出版	McKusick	誉为医学遗传学的圣经
1967	破译遗传密码	Holley, Khorana, Nirenberg	
1971	建立染色体 G 显带技术、定位第 1 个常染色体基因	Seabright, Donahue	
1975	建立染色体高分辨显带技术	Yunis	建立了微细胞遗传学
1976	重组 DNA 技术建立，提出遗传性肿瘤二次突变学说	Knudson	
1981	mtDNA 全序列测定；基因诊断技术	Anderson; Botstein	
1985	PCR 技术	Mullis, Saiki, Erlich	
1986	FISH 技术	Penkel	建立了分子细胞遗传学
1990	人类基因组计划(HGP)启动	Watson, Collins	誉为生物医学领域的阿波罗登月计划
1991	基因治疗进入临床试验	Anderson, Hott	
	发现线粒体遗传病	Wallace	
1994	人类基因组连锁图发表	Murray, Eissenbach, WhiteWard, Dausset	
1998	人类基因组物理图发表	Deloukas, Schuler, Gyapay, Beasley	
2000	人类基因组全序列工作框架图完成	美、英、日、法、德、中六国	
2001	人类基因组序列草图及初步分析结果完成	美、英、日、法、德、中六国	

第三节　我国医学遗传学的研究现状

从总体上看，虽然我国医学遗传学研究与国际先进水平有一定差距，但目前，这方面的研究正蓬勃开展，并已取得了不少的成就，在某些领域已经达到国际先进水平。

人类基因组研究是医学遗传学研究的热点。美国投资30亿美元用于人基因组作图与测序研究，我国国家“863”高新技术发展计划也在“八五”和“九五”期间重点资助了人类基因组研究。我国科学家于1999年9月加入国际人类基因组计划，承担人类3号染色体短臂上的大约30 Mb区域的测序任务，该区域约占人类基因组全部序列的1%。中国是参与这一研究计划的惟一发展中国家，正如国际人类基因组计划的“掌门人”Collins博士评价那样：“在这个划时代的里程碑上，已经重重地刻下了中国和中国人的名字。”

在人类基因组作图方面，复旦大学遗传研究所对人X染色体和17号染色体的部分空白区域进行了YAC排序和大尺度物理作图，填补了国际同类研究的空白，而且，作为整个人类基因组作图的一部分，其工作已经得到国内外同行的认可。湖南医科大学采用染色体显微切割方法分离特异区域探针，取得了成功。这项研究对于疾病相关基因的定位克隆具有重要价值。上海第二医科大学附属瑞金医院在揭示急性早幼粒细胞性白血病分子机制的基础上，克隆定位了位于11号染色体上的白血病相关基因，采用差异显示PCR方法克隆了2个全反式维甲酸诱导的新基因，目前正在进行大规模骨髓细胞表达序列标签(EST)的测定。

2000年底，中华民族基因组单核苷酸多态性(SNPs)系统目录的构建与研究，为我国在下一轮更加激烈的功能基因组研究的国际大竞争中打下了坚实的基础。

目前，我国科学家正在积极建立蛋白质组学的技术平台，建立高通量药物筛选平台和亚洲第一生物信息平台，以人类基因组为契机，计划从蛋白质组学到家禽基因组学，从生物芯片到干细胞，全方位切入生命科学的各个领域。

在基因诊断方面，我国已经能够对多种遗传病、病毒性疾病以及部分肿瘤进行基因诊断，特别是病毒性疾病的基因诊断已经在基层医院广泛开展。能够对β珠蛋白生成障碍性贫血(β地中海贫血)、亨廷顿(Huntington)舞蹈病、假性肥大性肌营养不良症等疾病患者、携带者和胎儿做出正确诊断。近年来，PCR技术在临床传染病的基因诊断中广泛应用，已成为重要的临床诊断方法。

国家“十五”重点攻关项目“人口出生缺陷干预工程”已全面启动，根据该项工程，在孕前、孕中、出生后3个阶段，均有相应检测方法发现胎儿或新生儿的异常，其中的基因芯片检测技术，通过对怀孕3个月以上的孕妇少量羊水的检测，就可判断胎儿是否患有遗传性疾病，进而针对不同情况，或终止妊娠，或及时进行治疗，以提高出生人口的质量。

基因治疗在理论上是根治遗传病的惟一方法，是当代医学的尖端项目，也是未来医学的发展方向之一，目前已成为生命科学领域的研究热点。为了使基因治疗能够安全、规范地进行，国家卫生部制定了人体细胞基因治疗质控要点。国内复旦大学遗传研究所甲型血友病基因治疗研究工作已开展十几年，并于1991年进行了世界首次乙型血友病基因治疗，取得了安全有效的结果，整体上达到世界领先水平。

目前，我国的基因治疗研究在高效、靶向性基因导入、外源基因导入后的有效调控等关键技术等方面实现了重大突破，基因治疗研究已从单基因遗传病扩大到多基因遗传病、病毒性疾病、心血管疾病、神经系统疾病、自身免疫疾病、内分泌疾病和肿瘤等。已被国家批准进行基因治疗的临床实验包括：凝血因子Ⅸ基因治疗乙型血友病、TK基因治疗人恶性胶质瘤、血管内皮细胞生长因子(VEGF)基因治疗胃癌、重组腺病毒－*p*53基因治疗头

颈部肿瘤等。

多基因遗传病的研究目前也已经成为国际医学遗传学研究的热点。我国作为多民族的人口大国，拥有其他国家不可多得的人类遗传资源，这是克隆疾病相关基因的关键。但是，人类遗传资源的研究利用具有时间紧迫性，我国的遗传病的流行病学调查不够深入，人类遗传资源还是未知数。随着独生子女的增多，人口流动和国外大公司的大举进入，人体遗传资源正在不断萎缩和流失。随着人类基因组研究的深入和相应技术的发展，只要有了若干重要的遗传病家系，就可以将疾病相关基因定位乃至克隆。如果在一些人口流动较小、交通不方便的不发达地区深入开展人体遗传资源的调查，可望挖掘到具有我国特色的遗传资源，为我国的分子遗传学研究提供良好素材，使我国能够在遗传病性疾病相关基因研究方面在国际上占有一席之地。由于家系资源在多基因遗传病研究中处于关键地位，我国丰富而独特的人体遗传资源使得我国在多基因遗传病研究中具有独特的优势；如果选择合适的病种，我国在多基因遗传病相关基因的研究的国际竞赛中可望取得成功。

第四节　医学遗传学在医学中的地位

医学遗传学的任务在于揭示各种遗传性疾病的遗传规律、发病机制、制定诊断和防治措施。随着医学的进步和人类生活水平、医疗水平的提高，早先严重危害人类生存和健康的传染性疾病已趋于绝迹或基本得到控制。与此相对，遗传病及与遗传因素密切相关的人类疾病所占比例日益突出，对人类健康的威胁日益严重。据统计，目前已发现的人类单基因遗传病达 6 000 多种；染色体畸变综合征在 100 种左右，加上异常核型已达近千种；多基因遗传病也在 100 种左右，多为常见病，如糖尿病、冠状动脉粥样硬化性心脏病(冠心病)、高血压病、精神分裂症等。据保守估计，有 1/5 ~ 1/4 的人患某种遗传病或与遗传有关的疾病。医学遗传学将研究这些疾病的发病机制，探索遗传因素、环境因素在发病中所起的作用，特别是对多基因遗传病中易感主基因的研究，对预防或减少这些常见病的发生有着重要意义。

肿瘤作为威胁人类生命、致人死亡的三大杀手之一，其发病也与遗传物质的改变有关，包括癌基因的激活和抑癌基因的失活。肿瘤遗传学对癌基因和抑癌基因的研究，是当前医学遗传学的热点，也是现代医学的迫切要求，一旦获得突破性进展，将是对现代医学的巨大贡献。

随着重组 DNA 技术的发展，许多遗传病、传染病已可以进行基因诊断、产前诊断，某些遗传病和恶性肿瘤等的基因治疗研究也正在积极进行，这些分子水平上的诊断和治疗的研究是医学遗传学为临床实践提供的重要手段，有着广阔的应用前景。

工农业生产的发展使环境污染日益严重；许多工业化学品或农药等均有诱变作用，从而增加了人类的遗传负荷。医学遗传学提供的检测突变的技术方法为降低遗传负荷、减少遗传病的发生起重要作用。

通过优生学的研究，减少或预防遗传病患儿的出生，提高后代的健康素质是医学遗传学的任务之一，也是现代医学的目标。

综上所述，可以看出，医学遗传学既是遗传学中发展最快的领域之一，同时也已经成

为现代医学极为重要的方面，是一个十分活跃的领域。医学遗传学的进展对推动医学的发展必将起到越来越重要的作用。

另外，医学遗传学在医学教育中，不仅与生物学、生物化学、微生物学与免疫学、病理学、药理学、组织学与胚胎学、卫生学等基础医学学科密切相关，而且已渗透到各临床学科之中。这说明，医学遗传学是医学教育中不可或缺的一门学科，作为一个医学生，要使自己成为一名合格的医务工作者，应当重视医学遗传学课程的学习。

第五节　遗传病概述

一、遗传病及其特征

（一）遗传病的概念

所谓遗传病（genetic disease），是指遗传物质结构或功能改变所导致的疾病。细胞中的遗传物质主要存在于细胞核内，少数存在于细胞质中的线粒体内，即 mtDNA 分子。不管是核内遗传物质的改变，还是线粒体内 mtDNA 分子的改变，都将引起遗传病。

通常人体内遗传物质及其控制的代谢方式与人体的周围环境保持平衡，遗传物质的缺陷或周围环境的显著改变，都能打破这种平衡，从而引起疾病。在不同疾病的病因中，遗传因素和环境因素所占比重各有不同，外伤、中毒、营养缺乏性疾病显然是环境因素引起的；另一些疾病则主要是遗传性的，如基因突变引起的半乳糖血症，染色体畸变引起的唐氏（Down）综合征，这些疾病只发生于有异常基因或有异常染色体的个体；还有一些疾病，其异常的遗传物质改变了个体代谢方式或代谢水平，但在一般情况下，仍可为人体所耐受，只有在接触特殊环境条件时才发病，如 G6pD 缺乏者在食用蚕豆或服用伯氨喹啉等药物后发生溶血危象。许多常见病，如高血压、糖尿病、消化性溃疡等的发病，既有遗传因素，又有环境因素，环境因素往往是诱因。因此，可以认为，任何疾病的发生都是环境因素与遗传因素相互作用的结果。

（二）遗传病的特征

1.遗传性　“遗传性”是指性状或疾病由亲代向子代传递的现象。大多数遗传病不同于传染病的水平传递，具有亲代向子代传递的特点。这是因为遗传病的发生是亲代生殖细胞或受精卵的遗传物质发生了改变所导致的，但不是每个遗传病的家系中都可观察到这一现象。有的患者是家系中的首例，即首次突变产生的病例；有些遗传病，比如，隐性遗传病，双亲不发病，子代可发病；而许多染色体畸变综合征患者活不到生育年龄或不育。在这些情况下，都看不到亲代向子代传递的现象。少数遗传病是体细胞内遗传物质的改变所引起。例如，急性放射病可导致体细胞遗传物质的损伤；肿瘤（非遗传性肿瘤）患者肿瘤细胞核的异常等，这些称为体细胞遗传病，通常不传给后代，故也无亲代向子代传递的现象。

2.先天性　“先天性”是指疾病生来就有的。例如，白化病是一种常染色体隐性遗传病，刚出生时就有“白化”表型。但是，不是所有的遗传病都是先天性的，有些遗传病致病基因的作用在出生后须经过漫长的生命过程才逐步表达。例如，甲型血友病患者出生后

无异常，一般在儿童早期才发病；再如，亨廷顿舞蹈病是一种典型的常染色体显性遗传病，往往要到35岁以后才发病。从另一方面来看，即使是先天性疾病（congenital disease）也并非全是遗传病。例如，母亲在孕期感染风疹病毒，可致胎儿发育异常；孕妇服用某些药物也可致胎儿发育异常。这些疾病都是在子宫内获得的，因此，是生来就有的，但却不是遗传病。

3.家族性 “家族性”是指疾病的发生呈现家族聚集现象，即在一个家庭中不止一个成员罹患，亲代和子代中均有患者。不少遗传病，特别是显性遗传病，常见家族聚集现象，但也有不少遗传病，特别是隐性遗传病和染色体病，无家族史，表现为散发性；在常染色体显性遗传伴外显不全的情况下，也表现出散发性。从另一方面看，即使是家族性疾病（familial disease），也不一定就是遗传病。例如，一个家族中某一成员患了某种传染病，如结核病或病毒性肝炎等，由于相同的生活环境条件，可以导致家族中其他成员也患同样的疾病；再如，夜盲症、甲状腺功能低下所致的痴呆症均有家族性，前者是由于饮食中缺乏维生素A所致，后者是由于碘摄入不足所致，都不属于遗传病。

二、遗传病的分类

根据遗传物质的结构和功能改变的不同，将遗传病分为4大类。

（一）单基因遗传病

人类体细胞中的染色体是成对的，染色体上的基因也是成对的，如果某种遗传病的发病涉及一对基因，由此导致的疾病称为单基因遗传病。单基因遗传病呈明显的孟德尔式遗传。目前已发现的单基因遗传病达6 000种以上。根据致病基因是显性还是隐性，是位于常染色体上还是X染色体或Y染色体上，单基因遗传病分为常染色体显性遗传病、常染色体隐性遗传病、X连锁显性遗传病、X连锁隐性遗传病和Y连锁遗传病。

线粒体中所含的mtDNA是独立于细胞核染色体外的遗传物质，也称为线粒体基因组。这些基因突变所导致的疾病，称线粒体基因病，随同线粒体传递，呈母系遗传。此类疾病也属于单基因遗传病，称线粒体遗传病。

（二）多基因遗传病

多基因遗传病是由多对基因与环境因素共同作用产生的遗传病。多基因遗传病的发生包括2种情况：一是由多对微效基因和环境因素共同作用引起；二是一对主基因和其他基因及环境因素共同作用所引起。由于这类疾病涉及的基因多，遗传机制复杂，所以，只有通过对大量患者的研究，才能确定遗传基因在多基因遗传病中的作用。

（三）染色体病

人类正常体细胞中具有二倍数染色体，包括22对常染色体和1对性染色体，每条染色体上有许多基因。如果在生殖细胞发生或受精卵早期发育过程中发生差错，导致染色体的数目或结构异常，就会产生整条染色体或染色体节段超过或少于二倍体数的个体，在这样的个体，由此所产生的疾病称为染色体病。由于发生染色体病时往往产生涉及许多基因的结构或数量的改变，故常表现出复杂的临床综合征，对个体的危害要远大于单基因遗传病或多基因遗传病。

(四)体细胞遗传病

体细胞内遗传物质改变所产生的疾病称为体细胞遗传病。这类遗传病一般不向后代传递,但随着细胞分裂增殖,可产生具有同样遗传物质改变的子细胞。例如,各种肿瘤的发病都涉及到特定组织细胞中的染色体或癌基因、抑癌基因的变化,故肿瘤属于体细胞遗传病。另外,某些先天性畸形也属体细胞遗传病。

三、遗传病的识别

前已述及,不同的遗传病遗传因素在发病中所起的作用大小不同,对于一种不明原因的疾病,要了解是不是遗传病,或者说是否有遗传因素参与,可采用如下一些方法。

(一)群体筛查

选定某一人群,采用简便、精确的方法对某种疑为遗传病的疾病进行普查。这种普查需在一般人群和特定人群(如患者亲属)中进行,通过患者亲属发病率与群体发病率的比较,从而确定某病是否与遗传有关。如果有关,则患者亲属的发病率应高于一般群体;而患者不同亲属之间,发病率还应表现为一级亲属大于二级亲属,二级亲属大于三级亲属,余类推。当发现某病亲属发病率较高时,为了排除同一家族成员的共同生活环境对发病的影响,应将血缘亲属与非血缘亲属加以比较,此时应见到血缘亲属发病率高于非血缘亲属。

(二)系谱分析

在初步确定某病为遗传病后,应搜集患者家族中全部成员的发病情况并绘成系谱,根据系谱,往往就能辨别是单基因遗传病,还是多基因遗传病,以及单基因遗传病中是何种遗传方式。

(三)双生子研究

同卵双生(monozygotic twin,MZ)的个体性别相同、遗传特性及表型特征也基本相同;异卵双生(dizygotic twin,DZ)的个体性别不一定相同,遗传特征及表型仅有某些相似,遗传基础与同胞相似。为了估计某种疾病或性状的发生中遗传因素所起作用的大小,可以对比 MZ 和 DZ 的发病一致性(concordance)的差异。发病一致性用发病一致率表示,即:

$$\text{发病一致率}(\%) = \frac{\text{同病双生子(同卵或异卵)对数}}{\text{总双生子(同卵或异卵)对数}} \times 100$$

如果这种一致性的差异越大,就表示这种病与遗传的关系越大;如果一致性的差异不显著,则表示遗传因素对发病所起作用较小或无作用。比如,在原发性癫痫,同卵双生的发病一致率为 60.1%,异卵双生发病的一致率为 9.4%,二者差异很大,这说明遗传基础在该病的发病中起相当重要的作用。

(四)种族差异比较

遗传流行病学调查表明,不同种族的个体,身体的外部性状如肤色、发型、发色、身材有差异,在血型、组织相容性抗原类型、血清型、同工酶谱等方面也存在着显著差异,这说明种族的差异具有遗传学基础。因此,如果某种疾病在不同种族中的发病率、发病年龄、性别和临床表现等方面有显著差异,则说明该病与遗传因素密切相关。当然,由于不同种

族生活的地理环境、气候条件、饮食习惯、社会经济状况等也会存在差异，故应严格排除这类环境因素的影响，最好将这种调查安排在不同种族居民混杂居住的地区进行。例如，在中国出生后移居美国的华侨，鼻咽癌的发病率比当地美国人高 30 多倍，这强烈提示鼻咽癌的发病有明显的遗传因素。

(五)疾病组分分析

对于比较复杂的疾病，特别是发病机制尚未完全阐明的疾病，在研究其是否与遗传因素有关时，可先将疾病"拆开"，即分解为若干环节(组分)，然后对各个组分进行单独的遗传学研究，如能确定某个或某些组分受遗传控制，则可认为该"组分"所在的疾病也受遗传控制。如冠心病是有复杂病因的疾病，高血脂症是其组分之一，已知家族性高胆固醇血症是常染色体显性遗传的，据此可以认为冠心病是受遗传控制的。

(六)关联分析

2 种性状或疾病非随机地同时出现，并且并非由于基因连锁所致，此称为关联。如 O 血型与十二指肠溃疡相关联；HLA - B_{27}与强直性脊椎炎相关联；HLA - B_8 与慢性活动性肝炎相关联等。如果已知某一性状或某一疾病决定于某基因位点的等位基因，即可以此作为遗传标记，检测另一性状或另一疾病是否与之关联；如果有关联，则表明后一性状或疾病也有遗传基础。关联的确切机制尚不明了，但作为遗传标记的基因直接或间接地参与了与之关联的疾病的发病，已是不争的事实。

(七)染色体分析

染色体病是遗传病的一大类，对多发性畸形、体格或智能发育不全的患者、孕早期反复流产的妇女，经过染色体检查、核型分析，可以确认是否为染色体病患者或携带者。

四、遗传病的危害

随着医学科学的进步、卫生水平的提高，急性传染病、流行病得到了有效的控制。人类的疾病谱已经改变，遗传病所占的比重越来越大，对人类健康的危害也越来越显著。

据有关资料，我国每年约出生 2 000 万人，其中约 1.3% 有严重的出生缺陷或先天畸形，遗传因素所致者占 80%，即每年大约有 20 万个新生儿因遗传因素产生严重的出生缺陷或先天畸形；自然流产中约有一半是由于染色体畸变所造成，每年约有 70 万例。有调查显示：我国城市儿童的死亡原因中，遗传病、先天畸形所占比例逐渐增大；与遗传有关的畸形包括先天性心脏病、大脑发育不全、消化道畸形、无脑儿、脊柱裂、脑积水等，约占先天畸形死亡总数的 90%，遗传病和先天畸形已成为儿童死亡的主要原因。智能发育不全的病例约有 1/3 是遗传病导致的，这是影响人口素质的重要因素。

从遗传病的发病率来看，有 20% ~ 25% 的人患遗传病或与遗传相关的疾病，其中染色体病约占 1%，单基因遗传病为 3% ~ 5%，多基因遗传病为 15% ~ 20%；体细胞遗传病中，恶性肿瘤对人类的危害最大，构成人类死亡第 1 位或第 2 位原因。另外，随着我国工业化进程的加快，人群正面临着环境污染的巨大威胁，由此将增高基因突变率，使人群的遗传负荷增高；增高的遗传负荷对人类的未来、后代的健康是不利的，这是一个严重的问题，是每一个医学遗传学工作者应该重视的问题。

小 结

医学遗传学是研究人类遗传病的发生机制、传递规律、诊断方法、治疗及预防措施的一门边缘学科。医学遗传学的研究领域和研究内容非常广泛，细胞遗传学、生化遗传学和分子遗传学是医学遗传学的主要组成部分，其他分支学科还包括群体遗传学、免疫遗传学、体细胞遗传学、肿瘤遗传学、药物遗传学等。目前，医学遗传学已发展成为由众多分支学科组成、涉及多门基础学科和临床学科的综合性学科。

医学遗传学的发展经历了缓慢、快速、飞跃发展3个阶段，自20世纪50年代起，随着生化分析技术和细胞遗传学技术的发展，医学遗传学进入了快速发展的时期；随着重组DNA技术和PCR技术的发展和应用，传统的医学遗传学发展进入到医学分子遗传学的水平，这两项技术不仅用于基因的结构、功能的研究及疾病相关基因的克隆，还用于基因文库的构建、基因探针的制备、基因表达和基因治疗的研究，正在对医学遗传学的各个领域产生重大影响，并取得了突飞猛进的发展。我国医学遗传学的研究虽然整体上与国际先进水平有较大差距，但研究工作正蓬勃开展，某些领域已接近国际水平。

医学遗传学是研究遗传病的科学，随着医学的进步、医疗水平的提高，人类的疾病谱已发生了重大变化，遗传病及与遗传相关疾病，包括冠心病、原发性高血压、糖尿病、肿瘤等，对人类健康的威胁日益严重。对这些疾病的发病机制、遗传因素和环境因素在发病中的作用的研究是医学遗传学的重大任务。因此，医学遗传学在现代医学中的地位日益突出，它是医学教育中不可缺少的一门学科。

遗传病是指遗传物质结构和功能改变所致的疾病。任何疾病的发生都是环境因素和遗传因素相互作用的结果。不同的疾病，二者所占的比例不同，大多数遗传病，遗传因素占主导地位，环境因素起着诱因作用。

遗传病可分为单基因遗传病、多基因遗传病、染色体病和体细胞遗传病4类。对于原因不明的疾病，要确认是否是遗传病，可采用群体筛查、系谱分析、双生子研究、种族差异比较、疾病组分分析、关联分析及染色体分析等方法。

遗传病对人类的危害越来越显著，积极开展遗传病的预防和治疗的研究工作、减少遗传病的发生、提高人口素质是医学遗传学工作者面临的艰巨任务。

（蔡绍京）

第二章　遗传的细胞基础

细胞是一切生物进行生命活动的基本结构单位和功能单位。虽然不同生物有机体或同一生物体的不同组织或器官的细胞大小、形态及功能彼此不同，但基本结构是相似的。真核细胞都是由细胞膜、细胞核和细胞质组成。细胞核在细胞的生命活动过程中处于极为重要的地位，是细胞内遗传信息储存、复制和转录的场所，也是细胞增殖、分化和调控的中心。

第一节　染色质与染色体

染色质(chromatin)是细胞核内能被碱性染料着色的物质，是遗传信息的载体，是在间期细胞核中呈伸展状态的 DNA 蛋白质纤维；染色体(chromosome)则是进入分裂期后高度螺旋化的 DNA 蛋白质纤维，是间期染色质结构紧密盘绕折叠的结果。因此，染色质和染色体实质上是同一物质在细胞周期的不同阶段，执行不同生理功能时呈现的 2 种不同存在形式。随着细胞从间期进入分裂期，染色质逐级螺旋，到分裂中期时，染色质高度凝缩为光镜下可见的棒状或杆状染色体；分裂期结束，染色体又解螺旋成为间期丝网状的染色质。

一、染色质的组成与结构

(一)染色质的化学组成

染色质是由核酸和蛋白质组成的核蛋白复合体，主要成分是 DNA 和组蛋白，还有非组蛋白及少量 RNA。

DNA 是染色质中储存遗传信息的生物大分子，是染色质中结构性质稳定、数量恒定的基本成分。不同动植物细胞染色质中 DNA 含量有所不同，但 DNA 含量并不随着生物体的复杂性而增加，许多植物细胞 DNA 量超过人类数倍。

组蛋白(histone，H)是染色质中富含精氨酸和赖氨酸的碱性蛋白，带正电荷，可与 DNA 紧密结合，对维持染色质结构和功能的完整性起关键作用。组蛋白是真核细胞特有的蛋白质，可抑制 DNA 的复制和转录。人类组蛋白包括 5 种：H_1、H_2A、H_2B、H_3 和 H_4，其中 H_1 是具有高度种属特异性的组蛋白，而 H_2A、H_2B、H_3 和 H_4 则是高度保守的核小体蛋白。

非组蛋白质(nonhistone protein，NHP)属酸性蛋白，是染色质中除组蛋白外的其他所有蛋白质的统称。非组蛋白质在细胞内的含量一般比组蛋白少，但种类繁多，功能各异。一些非组蛋白质为染色体的包装提供支架作用，而另一些则与基因的调控有关。目前已分离到 500 多种非组蛋白质。一般认为，非组蛋白质与组蛋白结合，能特异地解除组蛋白对 DNA 的抑制作用，促进 DNA 的复制和转录。

RNA 在染色质中的含量很低，这些 RNA 是染色质的正常组分，还是转录出的 RNA 的残余，目前尚未定论。

(二)染色质的结构及染色体的组装

人体细胞平均含有60亿碱基对的DNA,加起来有2 m多长,分布在46条染色体上。直径只有5~10 μm的细胞核如何能容纳2 m多长的DNA? 多年来,染色质的超微结构一直是细胞学研究的重大课题,其核心是DNA和蛋白质的结构关系以及这种关系的动态变化。从20世纪50年代末到70年代初,曾经提出过很多染色质的超微结构模型,由于缺乏充分根据,故这些模型没有得到公认。直到1973年,Olins关于染色质基本结构的研究取得了突破,确定了染色质的基本结构单位是核小体;在此基础上,Bak(1977年)提出了染色质的四级结构模型。

按照染色质的四级结构模型,染色质以核小体为基本结构,逐级包装压缩,经螺线管、超螺线管,最后包装成染色单体。

1.核小体(nucleosome)　核小体包括核心结构和连接部。核心结构由核心粒(core particle)和核心DNA两部分组成:核心粒是8个组蛋白分子(H_2A、H_2B、H_3和H_4各2分子)形成的八聚体,核心DNA长约140 bp,在核心粒外周缠绕1.75圈,形成直径为10 nm的"珍珠"状结构;连接相邻核心结构的DNA为连接DNA,长约60 bp;每段连接DNA的表面结合1分子H_1组蛋白,其长度大约是其他组蛋白的2倍,主要起稳定核小体结构的作用。连接DNA及其表面的H_1组蛋白合称为连接部。许多核小体彼此连接,形成一种串珠状结构,称为核小体串(图2-1),这就是染色质的一级结构。核小体核心结构的直径是10 nm,200 bp的核小体DNA长度约是70 nm(0.34 nm×200),由此推算,DNA包装成核小体,长度大约压缩了7倍。

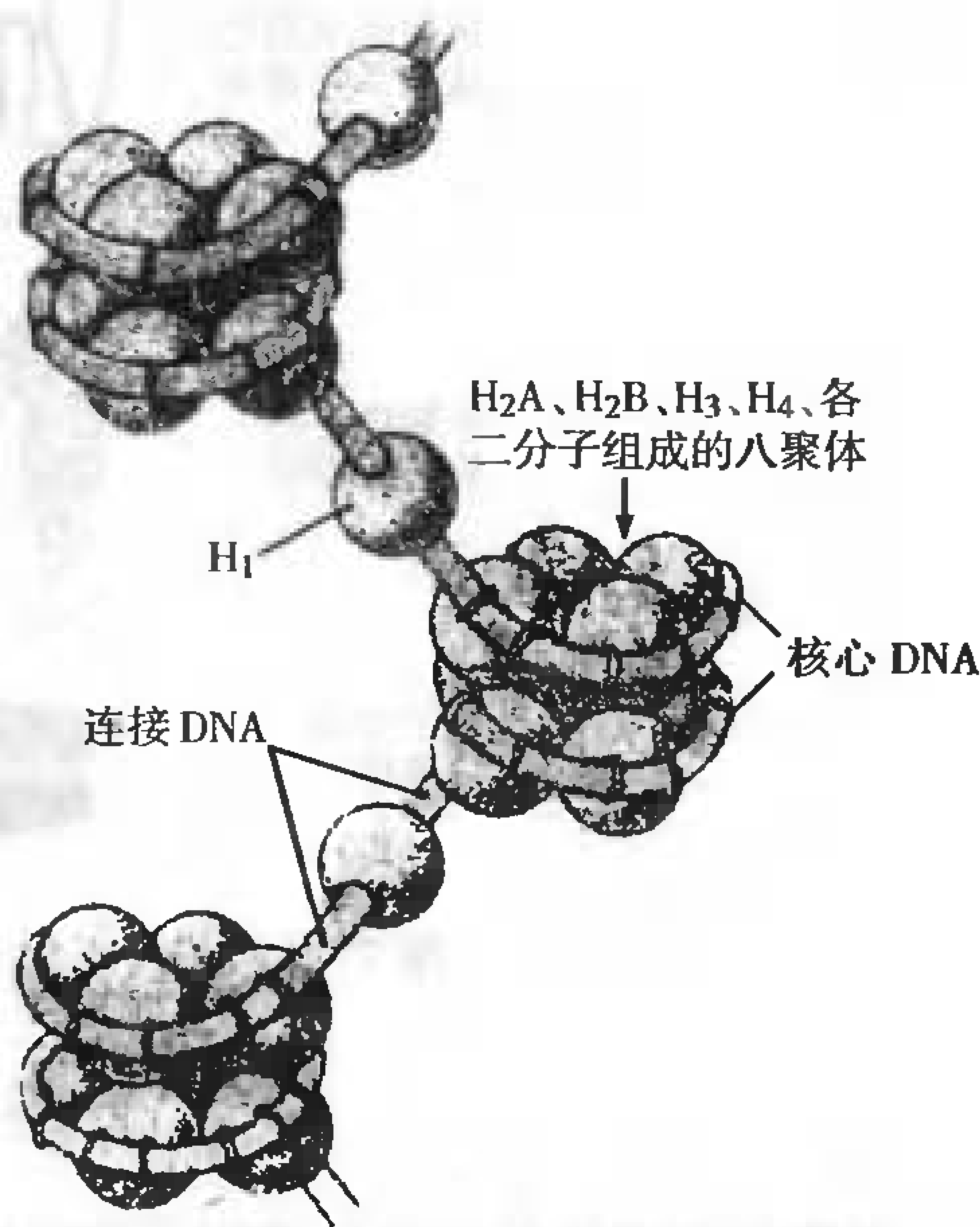


图2-1　核小体串

2.螺线管(solenoid)　核小体串螺旋化形成致密的、外径30 nm、内径10 nm、螺距11 nm的螺线管,又称染色质纤维。由于螺线管的每一圈含有6个核小体,因此,DNA长度又被压缩了6倍。螺线管是染色质的二级结构。

3.超螺线管(supersolenoid)　螺线管进一步螺旋化,形成直径为400 nm的超螺线管。在此过程中,DNA长度又被压缩了40倍。超螺线管是染色质的三级结构。

4.染色单体(chromatid)　超螺线管进一步螺旋化,形成直径为1~2 μm,长度为2~10 μm的染色单体,从超螺线管到染色单体大约压缩了5倍。由此推算,经多级螺旋至形成染色单体,DNA长度压缩了8 000~10 000倍(图2-2)。

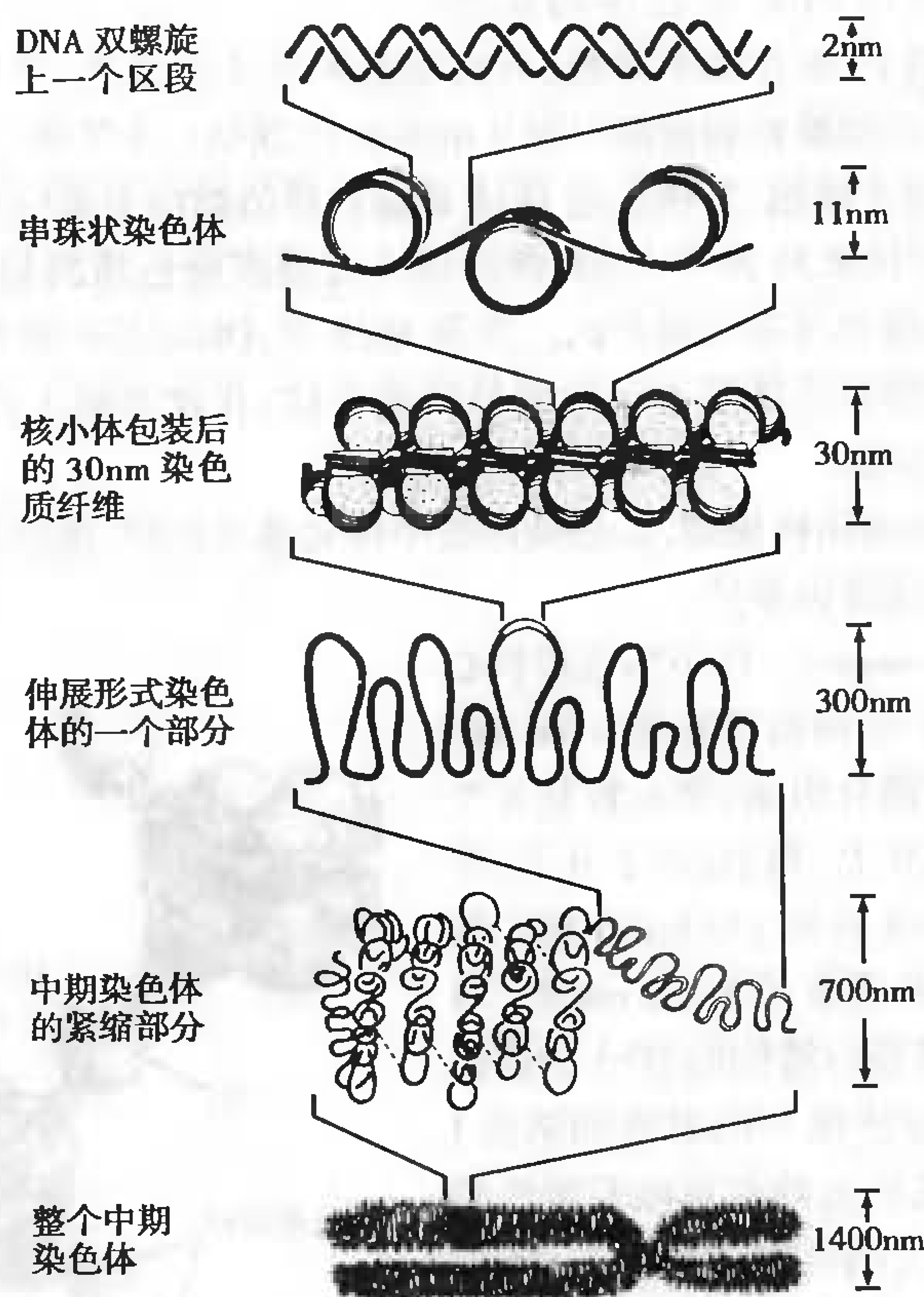


图 2-2　从 DNA 分子到中期染色单体的包装过程

二、染色质的类型

根据间期细胞核内染色质的形态特征、染色性能、折叠盘曲程度及功能，将染色质分为常染色质和异染色质 2 种类型。常染色质(euchromatin)是指染色质纤维折叠压缩程度低、处于伸展状态、着色浅的染色质。由于常染色质不易着色，折光性强，光镜下难以辨认；电镜下常染色质均匀分布于核内，核的中央相对较多。异染色质(heterochromatin)是指染色质纤维压缩程度高、处于聚缩状态、着色较深的染色质。异染色质通常位于细胞核边缘和核仁周围，构成核仁相随染色质的一部分。

常染色质与异染色质不是固定不变的，常随细胞的类型及细胞所处的生活周期和生理状态的不同而有差异。例如，细胞在一定的发育阶段，原来的常染色质凝缩，丧失基因转录活性，变为异染色质，这类异染色质称为兼性异染色质。如女性体细胞中的 X 染色质。另外，也可按功能状态不同将染色质分为活性染色质(active chromatin)和非活性染色质(inactive chromatin)。绝大多数细胞中具有转录活性的基因只占基因总数的 10%以下，其余 90%以上的基因是没有转录活性的，这些没有转录活性的基因，大部分以常染色质形式存在，但不转录，只有 10%～20%被包装成异染色质。因此，活性染色质是指具有转

录活性的常染色质,非活性染色质是指不进行转录的染色质,既有异染色质,也有部分常染色质。

第二节 细胞分裂与生殖

细胞的分裂增殖是生命得以延续的保证。高等生物的细胞增殖主要有 2 种方式:有丝分裂和减数分裂。

一、有丝分裂

有丝分裂(mitosis)是大多数生物体细胞增殖的主要方式,在有丝分裂过程中,遗传物质复制一次,细胞分裂一次,使母细胞内的遗传物质均等地分配到 2 个子细胞中,保证了遗传物质在传递过程中的相对稳定性。

一个个体从受精卵开始到个体发育成熟,直至衰老死亡,都在不断地进行着细胞的有丝分裂。细胞从前一次分裂结束开始生长到下一次分裂结束为止所经历的全过程称为细胞周期(cell cycle)。细胞周期包括间期和分裂期。

(一)间期

间期(interphase)是两次分裂期之间的时期,该期细胞在形态上虽然没有明显变化,但细胞内的生物合成却非常活跃。间期分为 DNA 合成前期(G_1 期)、DNA 合成期(S 期)及 DNA 合成后期(G_2 期)3 个时期。G_1 期是细胞生长的关键阶段,在此期间细胞进行许多重要的、特别是与启动 DNA 合成、进入 S 期有关的生化活动,并进行物质与能量的储备。G_1 期末存在一个限制点(restriction point,R 点)。根据细胞是否通过 R 点,可将细胞分为 3 种类型:①继续增殖细胞,这类细胞的分裂旺盛,对于组织的更新有着重要的意义,如上皮基底层细胞、部分骨髓细胞等;②暂不增殖细胞,也称为 G_0 细胞,这类细胞通常不分裂,但受到一定的刺激,如手术或外伤等,即可恢复分裂增殖能力,如肝、肾细胞及淋巴细胞等;③不再增殖细胞,或称终末分化细胞,这类细胞已丧失了分裂能力,如神经细胞、肌肉细胞及成熟的红细胞等。S 期是细胞周期中最关键的阶段,此期主要进行 DNA 复制,还合成参与染色体组装的组蛋白、非组蛋白质以及与 DNA 复制有关的酶。G_2 期主要是为分裂期细胞做物质和能量的准备,如合成构成纺锤体的微管蛋白、染色体凝集因子、及有丝分裂促进因子等蛋白质。

(二)分裂期

细胞经过 G_1 期、S 期和 G_2 期后,便进入分裂期(mitotic phase),即 M 期。在 M 期中,细胞核的分裂和细胞质的分裂在时间和空间上紧密配合,相互依赖,相互制约,构成一个复杂而连续的动态过程。根据细胞核的形态变化通常将分裂期分为 4 个时期:前期、中期、后期及末期(图 2-3)。

1.前期(prophase) 前期发生的主要变化包括:染色质凝集、核仁解体、核膜破裂,有丝分裂器形成。染色质纤维经螺旋化、折叠,最终形成棒状或杆状的染色体,每条染色体由 2 条染色单体构成。

2.中期(metaphase) 此期染色体达到最大程度的凝缩,着丝粒两侧的动粒分别与本

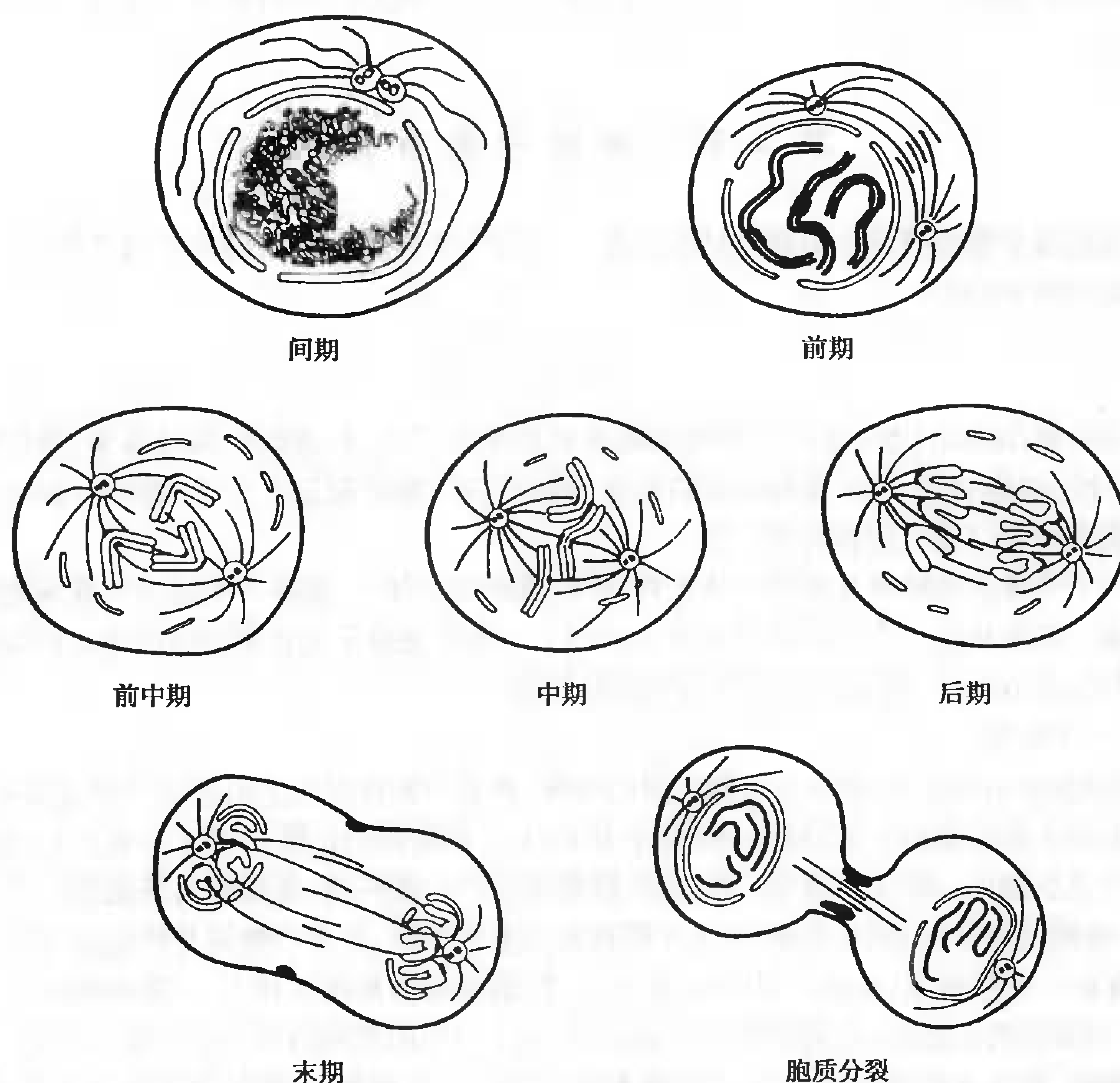


图 2-3　动物细胞的有丝分裂

侧的动粒微管相连。在微管牵引下，染色体向赤道面集中形成赤道板，呈现出最具典型特征的中期染色体形态结构。赤道板平面与纺锤体纵轴垂直。采用秋水仙素等抑制微管聚合的药物处理细胞，可阻断细胞分裂，使其停滞在中期，从而获得大量的中期分裂相。

3.后期(anaphase)　每条染色体的着丝粒纵裂，细胞内的染色体被平均分成两组，在动粒微管牵引下两组染色体分别向细胞的两极移动。后期染色体的向极运动，其动力来自动粒微管的不断缩短和极间微管的不断伸长。后期结束时，形态和数目相同的两组染色体到达细胞两极。

4.末期(telophase)　末期的主要变化是子细胞核形成和胞质分裂。到达两极的染色体解螺旋重新转变为染色质；消失的核仁在核仁组织区的位置上开始重新出现；分散在胞质中的核膜小泡与染色质表面相连，并相互融合，形成核膜。至此，2 个子细胞核形成。胞质分裂开始时，赤道板周围的细胞膜内陷，其下方的肌动蛋白和肌球蛋白大量聚集，形成收缩环(contractile ring)，继而形成分裂沟，后者不断加深，使细胞一分为二，完成胞质分裂。

二、减数分裂

减数分裂(meiosis)是一种特殊的有丝分裂,发生于有性生殖个体的配子形成过程中,故又称作成熟分裂(maturation division)(图 2-4)。减数分裂的主要特点是:DNA 复制 1 次,细胞连续分裂 2 次,1 个母细胞形成 4 个子细胞,子细胞中遗传物质减半。构成减数分裂的 2 次分裂称为减数分裂Ⅰ(meiosis Ⅰ)和减数分裂Ⅱ(meiosis Ⅱ)。

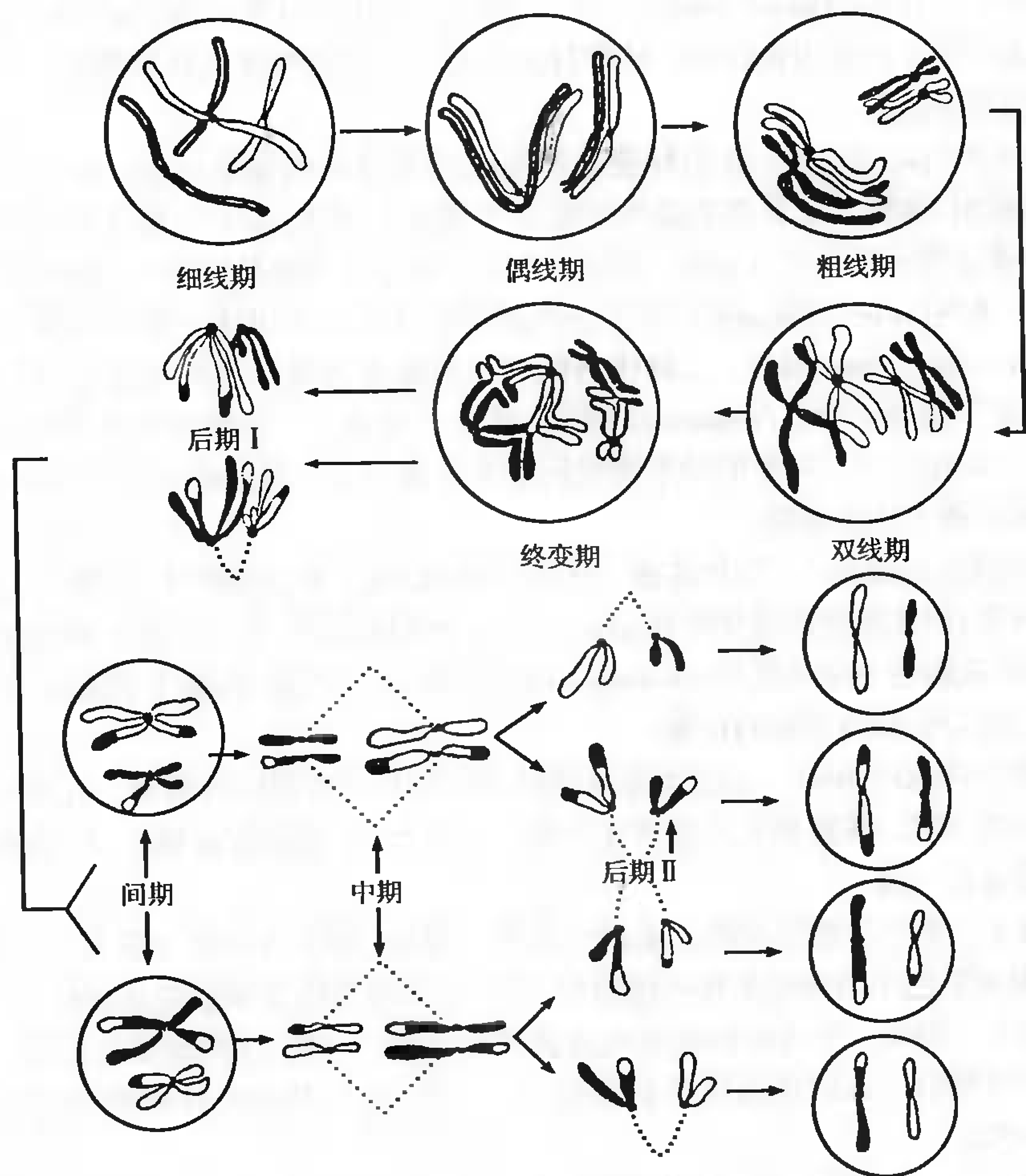


图 2-4　减数分裂的过程

(一)减数分裂Ⅰ

减数分裂Ⅰ可分为间期Ⅰ、前期Ⅰ、中期Ⅰ、后期Ⅰ和末期Ⅰ。

1. 间期Ⅰ　又称作减数分裂前间期(premeiosis interprase),与有丝分裂间期基本相同,也分为 G_1 期、S 期和 G_2 期。与有丝分裂间期的不同之处:一是 S 期明显延长,在有些物种,此期是有丝分裂间期的数十倍;二是 R 点存在于 G_2 期,如人的初级卵母细胞可长期停留在 G_2 期,只有接受性激素刺激后,才离开 G_2 期进入分裂期。

2. 前期Ⅰ　是减数分裂过程中历经时间最长,染色体变化最复杂的时期,通常将它分

成5个小的分期。

(1)细线期(leptotene) 染色质开始凝缩成细线状的染色体,但在光镜下不易区分,只见染色质丝盘绕折叠形成深染的染色粒(chromomere)。

(2)偶线期(zygotene) 最典型的特征是同源染色体配对。随机分布的同源染色体通过其一侧形成侧体彼此识别并相互配对,配对后2个侧体之间立即形成轴体结构,像拉链一样使同源染色体紧密相连。同源染色体配对的过程称为联会(synapsis),侧体与轴体总称为联会复合体(synaptonemal complex,SC),其化学成分均为蛋白质。联会的结果是每对同源染色体形成1个密切相伴的二价体(bivalent),人类46条染色体共形成23个二价体,但光镜下还难以区分。

(3)粗线期(pachytene) 非姐妹染色单体交换是该期的显著特征。染色体进一步螺旋化,缩短变粗,光镜下可见每条染色体包含2条染色单体,每个二价体则有4条染色单体,故二价体也称为四分体(tetrad)。四分体中,同属于1条染色体中的2条染色单体互称为姐妹染色单体(sister chromatid),而2条同源染色体的染色单体之间则互称为非姐妹染色单体(non - sister chromatid)。二价体内侧的非姐妹染色单体之间可发生多处点状的局部连接,光镜下呈现出交叉(chiasma)的现象,实际上是发生了非姐妹染色单体之间部分片段的交换(crossing over),交换的结果使得同源染色体中父源与母源的基因部分重组,这是生物遗传多样性产生的基础。

(4)双线期(diplotene) 染色体进一步缩短变粗,联会复合体解体,同源染色体相互排斥并逐渐分离,但在某些区域仍存在交叉。随着分离的进行,交叉点沿着染色体两臂向末端移动,这种现象称为交叉端化(chiasma terminalization)。因此,光镜下观察到的交叉点不一定是非姐妹染色单体交换的位置。

(5)终变期(diakinesis) 染色体变得更粗、更短,并向核的四周移动。此期结束时,纺锤体开始形成,核仁、核膜消失。随着交叉端化的进行,交叉点数目减少,但同源染色体的末端仍然紧靠在一起。

3.中期Ⅰ 各二价体排列在赤道面上形成赤道板,纺锤体微管与染色体一侧的动粒相连,2条同源染色体的动粒各朝向细胞的一极,但二价体仍借少数交叉相连。

4.后期Ⅰ 每个二价体或同源染色体彼此分开成二分体,非同源染色体随机组合,在纺锤体微管的牵引下分别移向细胞的两极。在人类,非同源染色体的随机组合,可形成2^{23}种组合方式。

5.末期Ⅰ 二分体到达细胞两极并逐渐解旋成为染色质;核膜、核仁重现;胞质分裂,形成2个子细胞。每个子细胞中都只有n个二分体,即染色体数目减少了一半,在人类是23条。

(二)减数分裂Ⅱ

1.间期Ⅱ 末期Ⅰ结束后,细胞进入短暂的间期Ⅱ,间期Ⅱ不进行DNA的复制,有的物种则没有间期Ⅱ,直接进入前期Ⅱ。间期Ⅱ以后的各期与有丝分裂过程基本相同。

2.前期Ⅱ 核仁、核膜消失,染色质重又折叠成染色体,每条染色体含2条染色单体(二分体)。

3.中期Ⅱ 各二分体排列在赤道面上形成赤道板,二分体上着丝粒两侧的动粒均与

本侧的动粒微管相连。

4.**后期Ⅱ**　各二分体上的着丝粒处纵裂，2条姐妹染色单体分离，形成2个单分体(monad)。在动粒微管的牵引下，单分体分别移向细胞的两极。

5.**末期Ⅱ**　到达细胞两极的单分体解旋成染色质，核仁、核膜重新出现，子细胞核重建，胞质也随之一分为二。结果每个子细胞中含有23条单分体，成为单倍体细胞。

经过连续2次的细胞分裂过程，人精原细胞或卵原细胞的2倍数染色体($2n=46$)变为成熟精子或卵细胞中单倍数染色体($n=23$)；精卵结合形成受精卵后，染色体数又恢复二倍数($2n=46$)，保证了亲代与子代遗传物质的相对稳定。

在后期Ⅰ，由于一对同源染色体中的2条染色体移向细胞的某一极是随机的，并且各对同源染色体的父源与母源染色体可以随机组合进入一个子细胞，所以，按理论推算，人类的23对染色体可有$2^{23}=8\,388\,608$种不同组合方式，即可形成8 388 608种不同染色体组成的配子；如果再考虑前期Ⅰ二价体的非姐妹染色单体间平均2.36次交换带来的重组变化，则更增加了配子中染色体组合的多样性。这一点正是有性生殖过程中表现的复杂遗传现象的基础。减数分裂不仅保证了物种及遗传性状的相对稳定，也是形成生物个体多样性的基础。

三、有丝分裂与减数分裂的比较

有丝分裂和减数分裂过程有许多相同之处。例如，都有细胞核与细胞质的周期性变化，都形成有丝分裂器，都存在染色体的螺旋化和去螺旋化等。但两者之间也存在明显区别。

有丝分裂过程中DNA复制1次，细胞分裂1次，1个亲代细胞形成2个遗传组成完全相同的子细胞，保证了遗传的稳定性；减数分裂过程，DNA复制1次，细胞连续分裂2次，1个亲代细胞形成4个遗传组成不同的子细胞，且遗传物质只有亲代细胞的一半，产生了遗传的多样性。

有丝分裂是体细胞的增殖方式，细胞中每条染色体都是独立的，不产生联会和交换；减数分裂则是配子形成特有的细胞分裂，减数分裂Ⅰ中发生同源染色体的配对及非姐妹染色单体之间的交换。

有丝分裂持续的时间一般为1～2 h，而减数分裂是个很长的过程，人的精细胞形成要持续70天左右，卵细胞形成则可持续十几年乃至数十年。

四、配子发生

有性生殖生物的精子与卵子的形成过程称为配子发生(gametogenesis)。

(一)精子的发生

精子是在睾丸的曲细精管中发生的。在6～7周龄的男性胚胎中就已出现了胚胎睾丸，但一直到男性青春期之前，睾丸中的精原细胞只进行有丝分裂，不产生精子。进入青春期，由于雄激素的诱导，精原细胞开始陆续分化为初级精母细胞，后者经过减数分裂产生精细胞，精细胞经过变态形成精子。

人类精子的发生过程可分为增殖期、生长期、成熟期和变形期4个阶段(图2－5)。

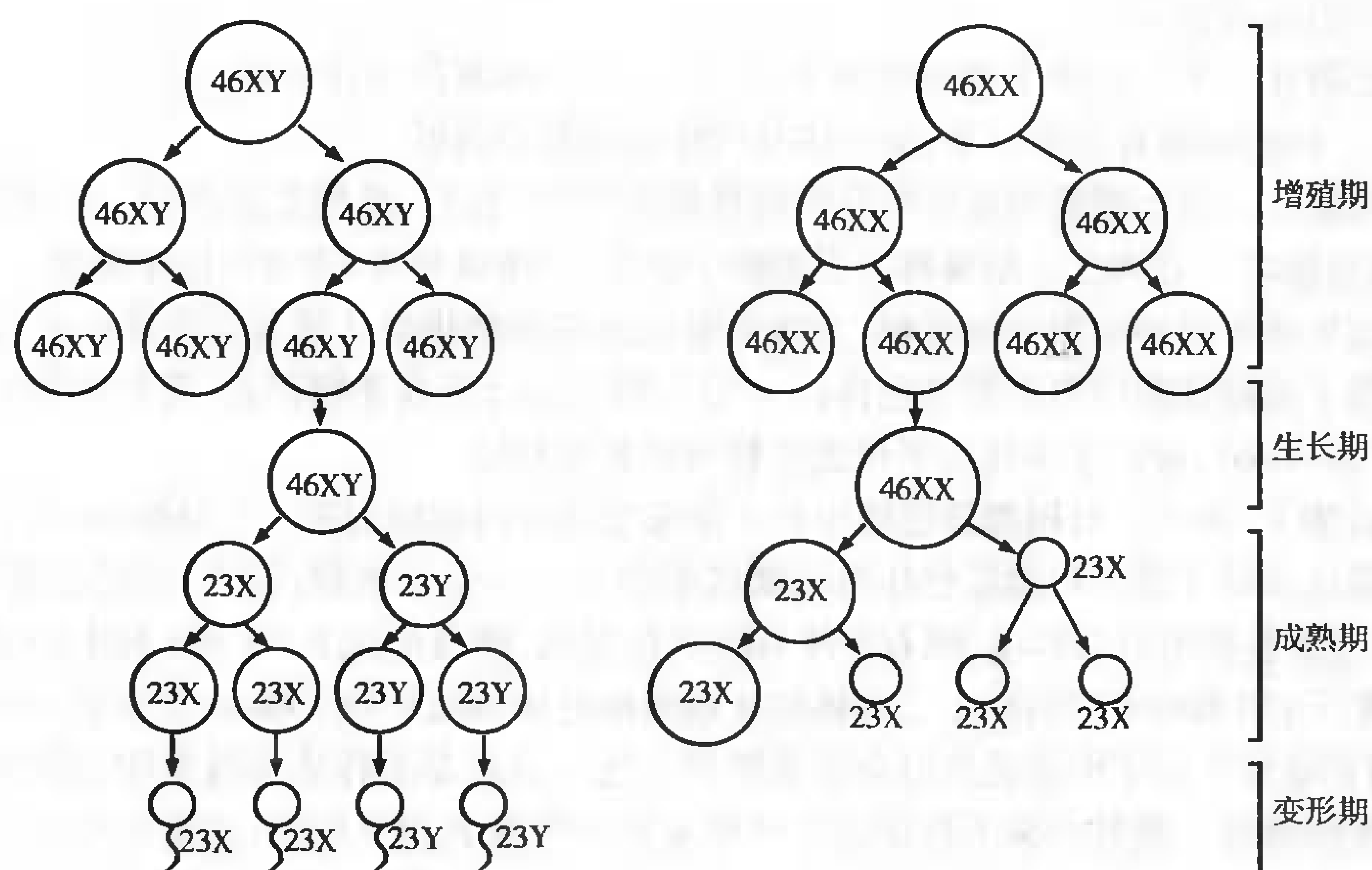


图 2-5　精子与卵子的发生过程

1.增殖期　睾丸曲细精管上皮中的精原细胞(spermatogonium)含有 46 条染色体,在增殖期经过有丝分裂,精原细胞数目增多。

2.生长期　精原细胞体积增大成为初级精母细胞(primary spermatocyte),此期细胞的数量和染色体数目不发生变化。

3.成熟期　初级精母细胞进行减数分裂。减数分裂Ⅰ后形成 2 个次级精母细胞(secondary spermatocyte),每个次级精母细胞中含有 23 条染色体,但每条染色体都有 2 条染色单体。1 个次级精母细胞经减数分裂Ⅱ形成 2 个精子细胞(spermatid),精细胞中仍含有 23 条染色体,但每条染色体只有 1 条染色单体。所以,经过减数分裂,1 个初级精母细胞形成 4 个精细胞,染色体数目从二倍数($2n$)变为单倍数(n)。

4.变形期　圆球形的精细胞逐渐分化为蝌蚪形精子。精子(sperm)分为头部、颈部和尾部,头部含细胞核,核的头端有由溶酶体转变而来的帽状顶体,内含蛋白酶和透明质酸酶。颈部含有一些巨大的线粒体,提供精子运动所需的能量。尾部的中轴部分是由微管组装的特殊结构,可保证尾部的摆动,使精子游动。

精子发生的 1 个周期为 70~72 天,男性自青春期起产生精子,一生中产生的精子总数可达 10^{12},即 1 万亿个。正常男性一次射精 3~4 ml,含 2 亿多个精子。

(二)卵子的发生

卵子在卵巢中产生,基本过程与精子发生相似(图 2-5),其不同之处主要有二:一是无变形期;二是初级卵母细胞的减数分裂过程呈间歇性。

1.增殖期　胚胎期女胎卵巢中卵原细胞(oogonium)经有丝分裂增殖,卵原细胞的染色体为 46 条。

2.生长期　在胚胎发育的第 3 个月,卵原细胞进入生长期,卵原细胞体积显著增大,

在胚胎发育的第 6 个月左右，所有卵原细胞均分化发育成初级卵母细胞(primary oocyte)。其细胞质中积累了大量的卵黄、RNA 和蛋白质等营养物质。初级卵母细胞的染色体仍为 46 条，每个初级卵母细胞被一层卵泡细胞所包围，构成原始卵泡(primordial follicle)。

3.成熟期　初级卵母细胞形成后便进入成熟期，开始进行减数分裂，但停留在前期Ⅰ的双线期，直到胎儿出生。进入青春期后，正常情况下，每个月都有初级卵母细胞继续双线期以后的细胞分裂。减数分裂Ⅰ形成 1 个次级卵母细胞(secondarg oocyte)和 1 个第一极体(first polar body)。次级卵母细胞迅速进入减数分裂Ⅱ，但停止在中期Ⅱ。含有该次级卵母细胞的卵泡排入输卵管，如果在 12 h 内与精子受精，则完成减数分裂，形成 1 个成熟的卵细胞并释放出 1 个第二极体(second polar body)；否则，次级卵母细胞退化消失。同时，第一极体也分裂为 2 个第二极体。一个初级卵母细胞经过减数分裂形成 1 个卵细胞和 3 个极体，它们都具有单倍数染色体($n=23$)，3 个第二极体均退化。

女性胚胎 6 个月时，卵巢中约有 700 万个原始卵泡；随着胚胎发育，原始卵泡退化增多，出生时尚有 100 万个；到青春期仅存 4 万个。正常女性 13～15 岁进入青春期，绝经年龄一般是 45～50 岁，这就是说，卵泡中初级卵母细胞到卵细胞的形成过程至少十几年，最长可达 50 年。女性一生中只有 400～500 个初级卵母细胞发育成熟形成卵细胞，仅占青春期卵巢中初级卵母细胞数(原始卵泡数)的 1%。

五、人类的性别决定

人类性染色体，即 X 染色体和 Y 染色体决定人类的性别。男性的性染色体组成是 XY，女性的性染色体组成为 XX，这种性别决定方式为 XY 型。

如前所述，人类的精子和卵子是通过初级精母细胞和初级卵母细胞减数分裂产生的，因此，男性可产生含有 X 染色体和含 Y 染色体的精子——X 型精子和 Y 型精子，2 种精子的数目相等；而女性只产生含 X 染色体的卵子。如果 X 型精子与卵子结合，就形成含有 XX 性染色体的受精卵，将来发育成女性个体；如果 Y 型精子与卵子结合，则形成含 XY 性染色体的受精卵，将来发育成男性个体。在自然状态下，2 种类型的精子与卵子的结合是随机的，因此，人类的性别比大致为 1:1。

很显然，人类性别实际上是由参与受精的精子中带有 X 染色体或 Y 染色体决定的，而 X 染色体和 Y 染色体在人类性别决定中的作用并不相等。一般情况下，胚胎的原始生殖组织在无雄性因子的作用下有向女性分化的倾向，即人类的性别决定于 Y 染色体，Y 染色体在性别决定中起关键作用，有 Y 染色体存在就可形成睾丸，无 Y 染色体存在则形成卵巢。例如，45,X 和 48,XXXX 患者呈女性表型，47,XXY 和48,XXXY患者则表现为男性。

Y 染色体的存在之所以使个体发育为男性，是因为 Y 染色体短臂上有一个与睾丸分化有关的基因，称为睾丸决定因子(testis－determining factor，TDF)基因。1990 年，Sinclair 等在 Y 染色体(Yp11.2)上发现了性别决定基因(sex－determining region of Y chromosome，SRY)，认为它是 TDF 的最佳候选基因，并证实它与人类性别决定密切相关。1993 年，Hua 等采用逆向遗传学方法对 SRY 进行了鉴定，确认长约 35 kb，无内含子，转录单位全长 1.1 kb。目前认为，人类的 Y 染色体由 2 个具有不同功能的区域组成：一是拟常染色质区(pseudo－autosomal region，PAR)，位于 Y 染色体短臂远端，X 染色体短臂上有与之对应的

同源区(图2-6)。在这一区域,X染色体和Y染色体间可发生配对和交换,此区内的交换并不表现典型的性连锁关系。另一区域称为Y特异区,位于PAR的近侧端,正常减数分裂过程中不与X染色体进行重组,SRY基因就位于这一区域,该区域对性腺分化及男性生育能力都是必需的。在减数分裂过程中,如果发生了Y特异区与X染色体之间的交换,就会出现XX男性或XY女性个体(图2-6)。

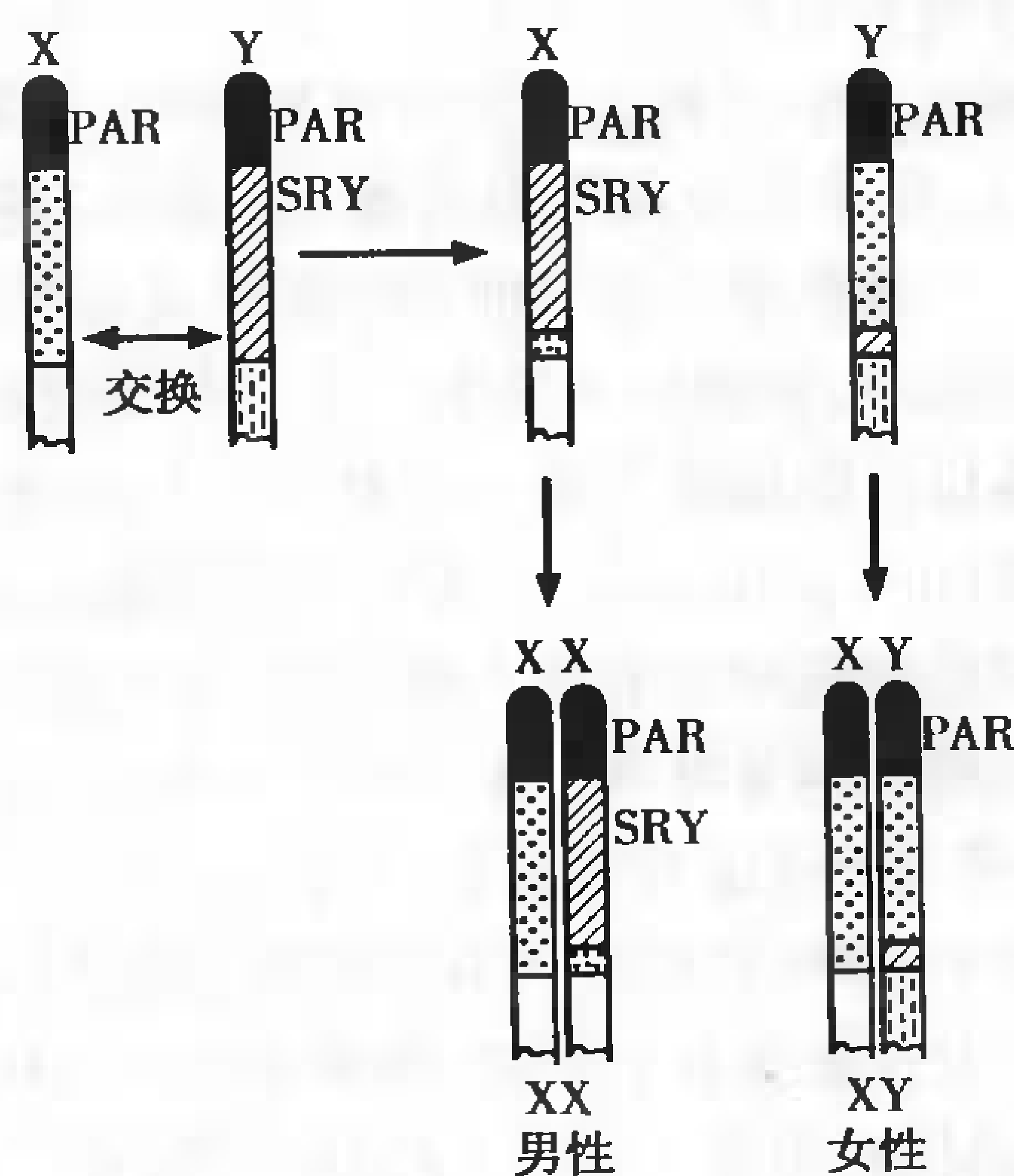


图2-6　性染色体的功能区与性别决定

尽管有越来越多的证据说明,SRY与人类性别决定密切相关,但却不能肯定SRY就是决定睾丸发育的惟一基因。例如,有多个X染色体的男性,其雄激素水平低下并有隐睾出现;常染色体三体型患者也常伴有睾丸异常;许多其他染色体异常导致的精子发育受损比相同情况下卵子发育受损程度要严重得多等。这些事实都说明,在性别决定中,除SRY基因外,还有其他基因的协同作用。

第三节　人类染色体

对人类染色体的研究早在19世纪就已开始,但直到1956年Tjio和Levan才确定了人类体细胞染色体的数目是46条。1959年,Down综合征、Turner综合征及Klinefelter综合征的染色体异常的发现,开创了细胞遗传学和临床疾病结合的新领域——临床细胞遗传学。20世纪,70年代,相继出现了多种染色体显带技术,提高了染色体分析的精确性。

一、人类染色体的形态结构

在细胞周期中,通常只在有丝分裂中期可以清楚地看到具有典型结构的染色体。每条染色体均含有2条姊妹染色单体,2条姊妹染色单体仅在着丝粒处相连,着丝粒区富含重复序列DNA构成的异染色质,浅染并内缢,故着丝粒区也称主缢痕(primary constriction)。着丝粒的外侧为动粒,是纺锤体微管附着处,在细胞分裂中与染色体的运动密切相关,失去着丝粒的染色体片段通常不能在分裂后期向两极移动而丢失。着丝粒将染色体沿纵轴分为两部分,较长的称为长臂(q),较短的称为短臂(p)。图2-7为中期染色体的模式图。

根据着丝粒的位置,将人类染色体分为3类:①中着丝粒染色体(metacentric chromosome),着丝粒位于染色体纵轴的1/2~5/8处,两臂长短相近;②亚中着丝粒染色体(submetacentric chromosome),着丝粒明显偏离中央,位于染色体纵轴的5/8~7/8处;③近端着丝粒染色体(acrocentric chromosome),着丝粒位于染色体纵轴的7/8至末端之间。

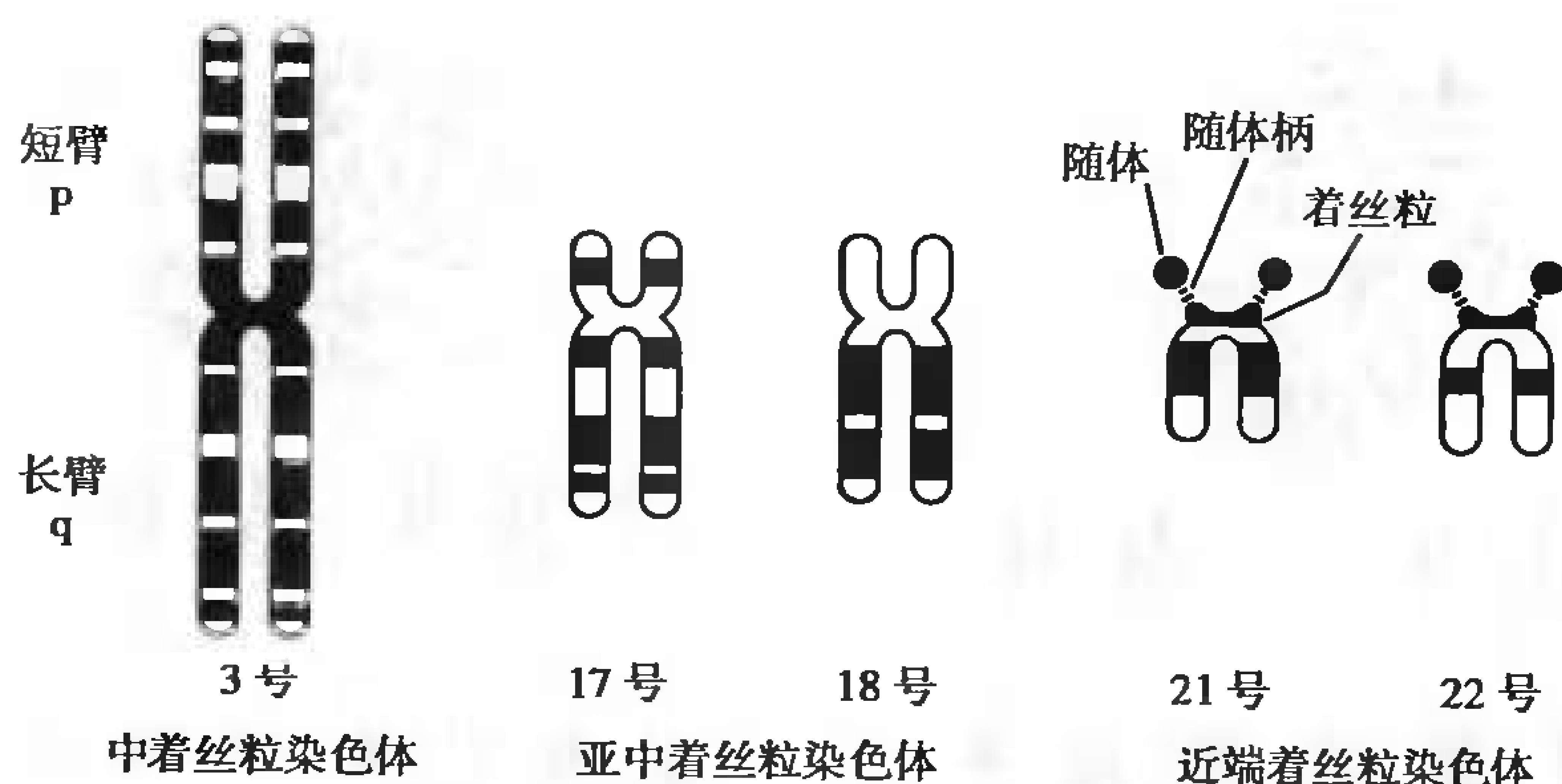


图2-7　人类染色体的形态结构

有些染色体的非着丝粒区也能见到浅染内缢的节段,称为次缢痕(secondary constriction)。人类近端着丝粒染色体(D组和G组)短臂上的次缢痕区(也称随体柄)是rRNA基因所在之处,转录的rRNA构成核仁主要成分,特殊的银染法可将其染成黑色。在该区的远侧可见到染色较深的球形小体,称为随体(satellite)。人类的5对近端着丝粒染色体短臂上均有随体存在,但也可能由于制片过程的影响造成随体丢失。另外,1、9和16号染色体长臂上也存在次缢痕。

在染色单体的两端,各有1个端粒(telomere)。端粒在维持染色体的稳定性和完整性方面起重要作用。缺失端粒的染色体,其末端彼此会黏合,形成易位或倒位等畸变。人类染色体端粒DNA含有(TTAGGG)n重复序列,长度5~20 kb。近年来的研究揭示体细胞染色体端粒的重复DNA序列长度的缩短与细胞的衰老和肿瘤的发生有关。

二、人类正常核型

(一)常规核型分析

核型(karyotype)是指一个有丝分裂中期的体细胞内的全部染色体按一定形式排列起来所构成的图形(图2-8)。按照Denver体制规定的分组排序原则进行剪贴配对的过程称为核型分析(karyotype analysis)。正常核型的记录的格式是:染色体总数,性染色体构成,二者间用逗号隔开。例如,46,XX代表正常女性核型;46,XY代表正常男性核型。

Denver体制是1960年在美国丹佛(Denver)召开的第1届国际细胞遗传学会议上确定的人类染色体命名标准体制,Denver体制将人体细胞的46条染色体按其大小和着丝粒位置分为A、B、C、D、E、F、G 7个组,X和Y染色体分别归入C组和G组,各组的基本特征如下。

1. A组(1~3号染色体)　最大,在非显带染色体标本上彼此能够区分。1号染色体与3号染色体均为中着丝粒染色体,1号染色体长臂近着丝粒处常见次缢痕。3号染色体比1号染色体约短20%,2号染色体为亚中着丝粒染色体。

2. B组(4号和5号染色体)　为较大的亚中着丝粒染色体,4号染色体略大于5号染色体。

3. C组(6~12号染色体及X染色体)　为中等大小的亚中着丝粒染色体。其中,6、

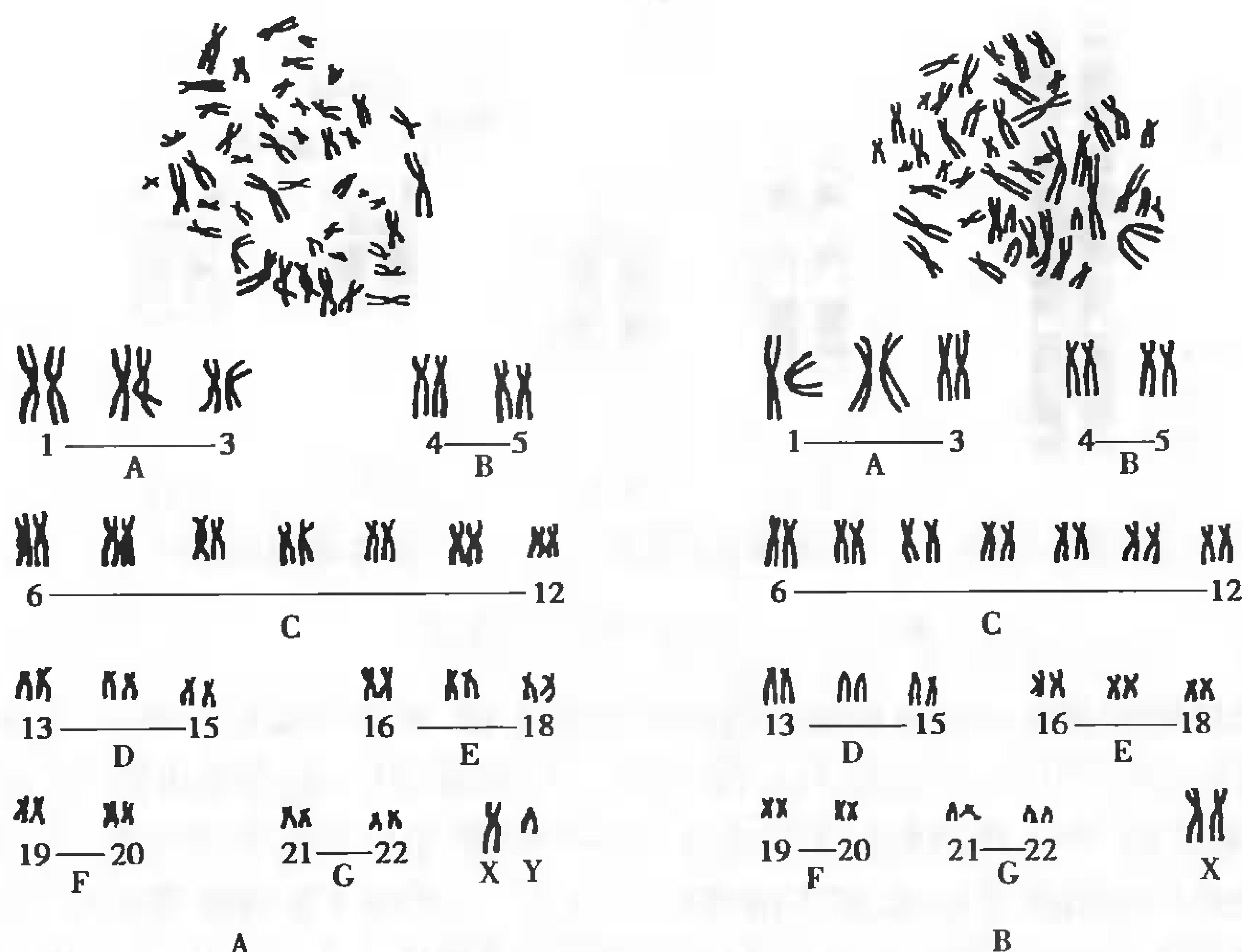


图 2-8　正常男性和正常女性的核型

7、11 号染色体及 X 染色体长短臂的差距较 8、9、10、12 号染色体小，9 号染色体长臂近侧常见次缢痕；X 染色体的大小介于 7 号染色体与 8 染色体号之间，长短臂的差距较 7 号染色体小。

4. D 组(13～15 号染色体)　为中等大小的近端着丝粒染色体。理论上这 3 对染色体都有随体，但由于制片过程中随体易丢失，故有时见不到随体。各号染色体鉴别困难。

5. E 组(16～18 号染色体)　16 号染色体为中着丝粒染色体，大约 10%个体 16 号染色体的长臂近着丝粒处有次缢痕。17 号和 18 号染色体为亚中着丝粒染色体。

6. F 组(19 和 20 号染色体)　是最小的中着丝粒染色体，19 号染色体比 20 号染色体略大。

7. G 组(21、22 及 Y 染色体)　是最小的近端着丝粒染色体。21、22 号染色体短臂末端均有随体，22 号染色体大于 21 号染色体。Y 染色体无随体，比 21 和 22 号染色体大，两长臂常平行伸展，长臂中部有时可见次缢痕。

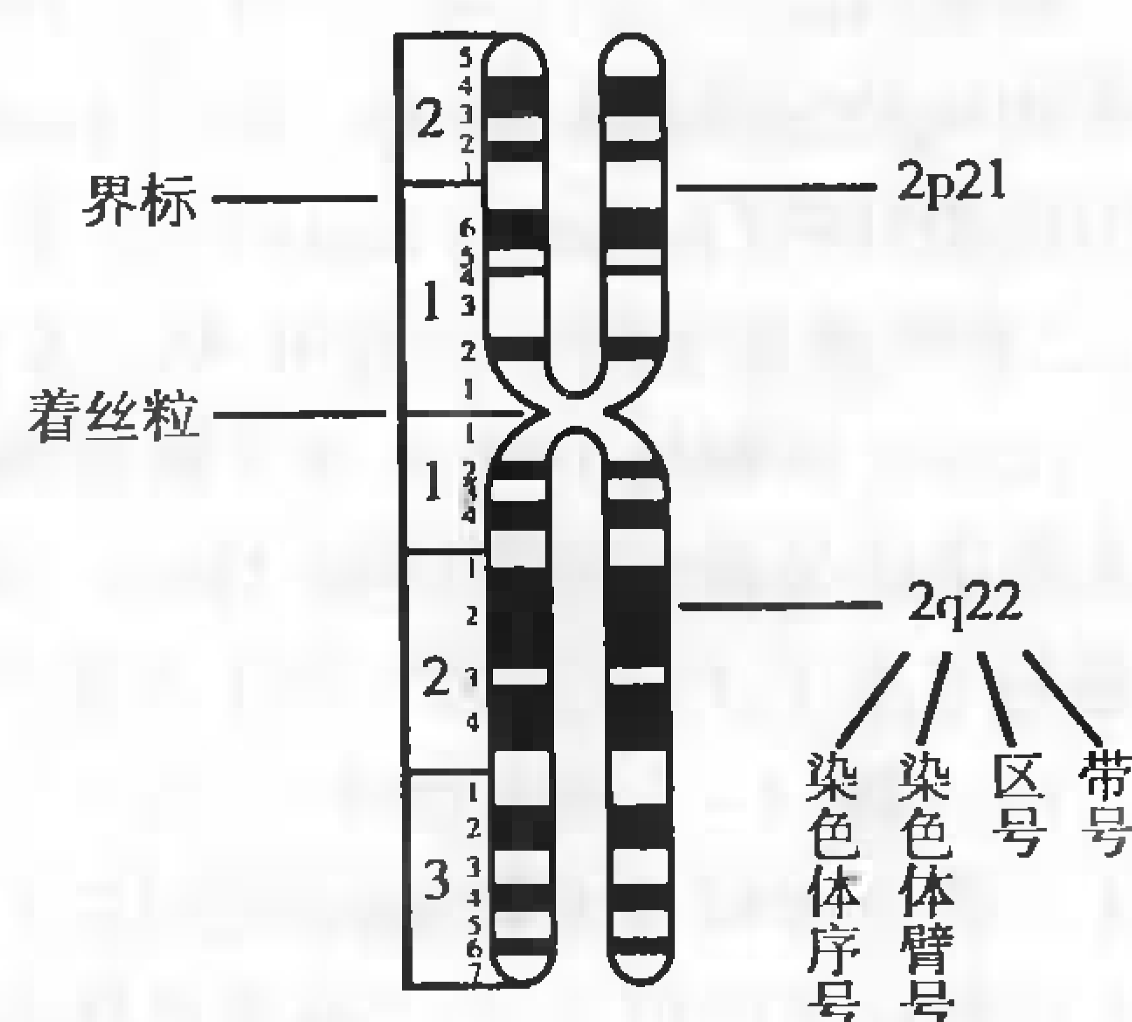


图 2-9　显带染色体模式图

(二)显带核型分析

1971 年，在巴黎召开的第 4 届国际人类细胞遗传学会议制定了显带染色体模式图(图 2-9)，提出了区分每个染色体区带的标准体系及记录规则。按此标准及规则，记述染色体中某一特定带时需写明 4 方面的内容：①染色体号，1～22 号染色体及 X 或 Y 性染色体；②臂的符号，长

臂 q 或短臂 p;③区的序号,“区”为位于染色体臂上两相邻界标之间的区域。“界标”是染色体上恒定的、有显著形态特征的带,它包括染色体两臂的末端、着丝粒以及臂上某些显著的带。区的序号是从着丝粒部位向两臂远端依次编号;④带的序号,“带”是染色体上宽窄各异、明暗相间的横纹。每一条染色体都应看做是由一系列连续的带构成的,即没有非带区。每条带因其着色的深浅而清楚地与相邻带相区分。带的序号也是从着丝粒部位向两臂远端依次编号。作为界标的带属于该界标以远区的第 1 条带,着丝粒分属长臂、短臂的第 1 区第 1 带。以上 4 个内容按顺序连写,不加标点。例如,2q21 表示第 2 号染色体长臂 2 区 1 带,1p31 表示第 1 号染色体短臂 3 区 1 带。

常规 G 带(图 2-10)的标准带型是 320 条带(单倍体染色体组的总带数),20 世纪 70 年代中、后期,随着细胞同步化方法的应用和显带技术的提高,产生了高分辨显带技术(high resolution banding technique),所谓高分辨显带主要是对晚前期、早中期或更早时期的染色体显带,可获得 550~850 条带,甚至上千条带纹,这种带纹数更丰富的染色体称为高分辨染色体。高分辨染色体带纹数的增多是由于对常规中期染色体显带中的某些带逐级细分所致,逐级细分产生的带称为亚带和次亚带。高分辨显带技术的应用,发现和证实了一些常规带型分析发现不了的细微的染色体畸变。所以高分辨显带技术对临床染色体病的诊断和肿瘤的研究具有重要意义。

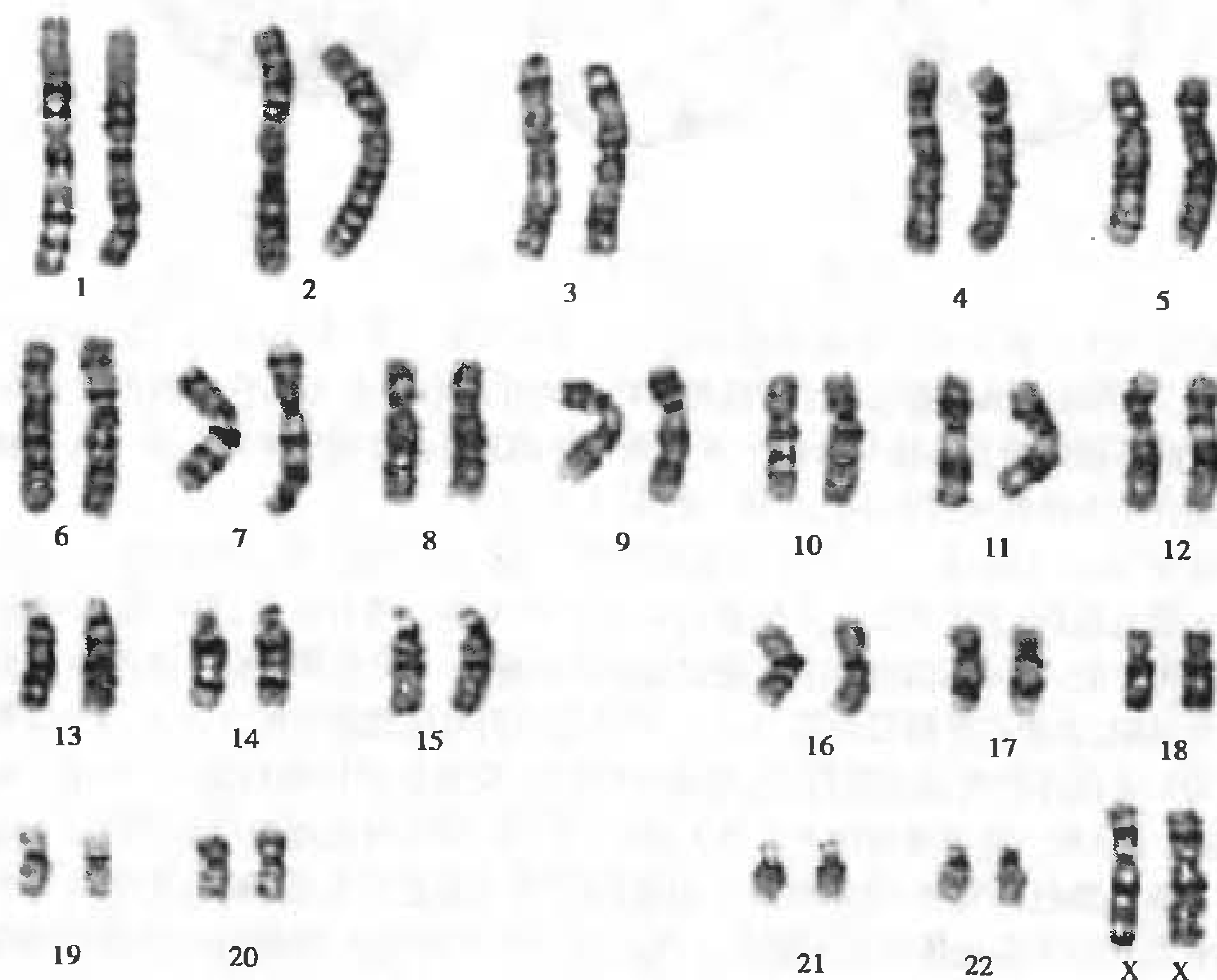


图 2-10 正常女性 G 显带核型

高分辨显带染色体的命名遵照人类细胞遗传学命名的国际体制(ISCN 1978,1981,1985):亚带书写在带号后面,以小数点相隔,编号原则也是从近侧向远侧依次编号,例如

1q42 可分为 3 个亚带，分别是 1q42.1、1q42.2 和 1q42.3；次亚带直接书写在亚带后，不加标点，如 1q42.1 再分为 3 个次亚带时，则分别写作 1q42.11、1q42.12 和 1q42.13。

三、性染色质

性染色质(sex chromatin)是间期细胞核中性染色体的异染色质所显示出来的一种特殊结构，人类性染色体有 X 和 Y 染色体，所以性染色质包括 X 染色质和 Y 染色质 2 种。

(一)X 染色质

1949 年，Barr 等在雌猫神经细胞间期核中发现了一种浓缩小体，直径约 1 μm，但雄猫神经细胞核中却没有这种结构(图 2－11A)。进一步的研究发现，除猫之外，其他雌性哺乳动物神经细胞及其他细胞间期核中也同样具有这种显示性别差异的结构。如大部分正常女性的上皮细胞、成熟中性粒细胞、成纤维细胞及口腔黏膜细胞间期核也有这种特征性的结构，而男性没有。这种与性别有关的结构称为 X 染色质(X chromatin)或 X 小体，又叫 Barr 小体。

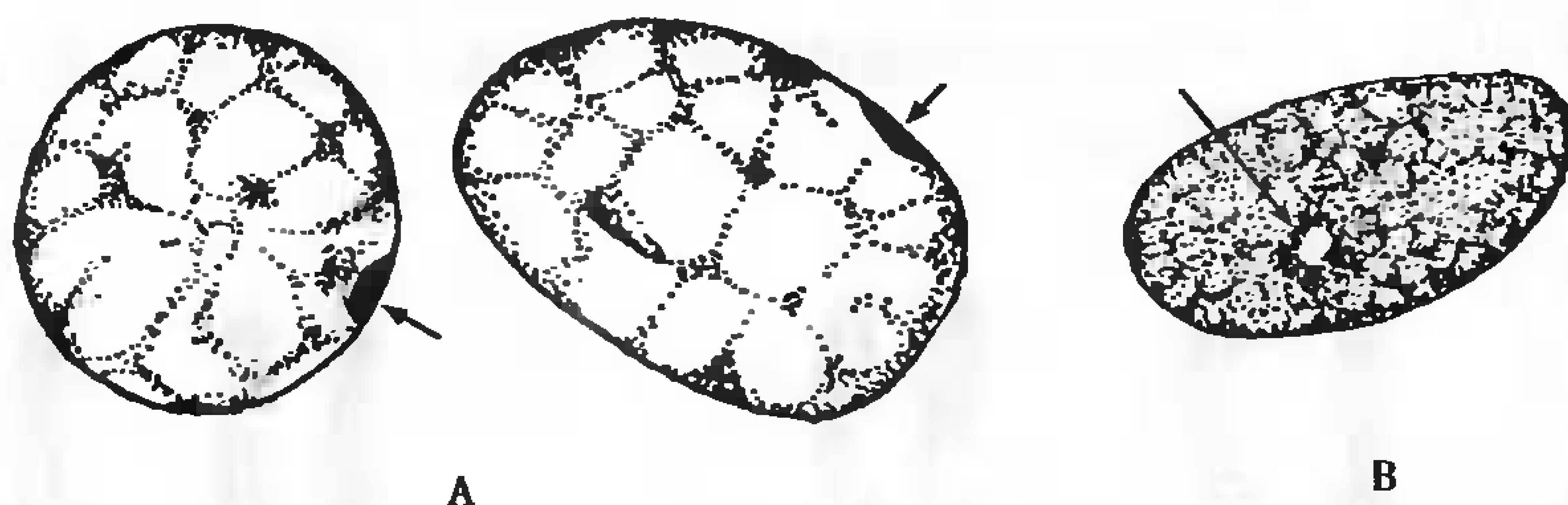


图 2－11　X 染色质和 Y 染色质

为什么人类正常女性间期细胞核中有一个紧贴核膜内缘、大小约 1 μm 的椭圆形深染小体，而正常男性间期细胞核中没有这种小体？为什么有 2 条 X 染色体的正常女性 X 染色体连锁的基因产物并不比只含有 1 条 X 染色体的男性多 1 倍？为什么某一 X 连锁的突变基因纯合子女性的病情并不比半合子的男性严重？

1961 年，Mary Lyon 提出的 X 染色体失活假说，即 Lyon 假说，较好地解释了上述问题。

Lyon 假说包括：①正常女性体细胞内仅有 1 条 X 染色体有活性，另一条 X 染色体在遗传上是失活的，在间期细胞核中螺旋化而呈异固缩的 X 染色质。②X 染色体的失活开始于胚胎早期，大约在受精后的第 16 天。在此之前所有体细胞中的 2 条 X 染色体都具有活性。③X 染色体的失活是随机的，即异固缩的 X 染色体可以来自父亲，也可以来自母亲。但是，一旦某一特定细胞内的 1 条 X 染色体失活，那么由此细胞增殖的所有子细胞也总是这 1 条 X 染色体失活，即如果某一细胞的父源 X 染色体失活，则由其分裂而形成的子细胞中失活的 X 染色体也是父源的。因此，X 染色体的失活是随机的，但却是恒定的。

间期细胞核内 X 染色质数目总比 X 染色体数目少 1，即 X 染色质数＝X 染色体数－1。例如，核型为 46，XX 的正常女性有 1 个 X 染色质；47，XXX 的超雌综合征女性则有 2 个 X 染色质；而核型为 46，XY 的正常男性及核型为 45，X 的 Turner 综合征女性细胞中 X 染色质数为 0。也就是说，当某一个体具有额外的 X 染色体时，该染色体也是失活的，这

种个体细胞中也只有1条X染色体具有遗传活性；对于正常男性及具有1条X染色体的女性，单条X染色体并不发生异固缩的现象，在任何时候都有活性，所以细胞中不存在X染色质。

正是由于正常女性体细胞中的1条X染色体发生了异固缩，失去了转录活性，这样就保证了男女性个体X染色体上的基因产物在数量上基本一致，这称为X染色体的剂量补偿(dosage compensation)。需要指出的是，失活的X染色体上一部分基因仍保持着一定活性，因此，X染色体数目异常的个体在表型上还是有别于正常个体。例如，47,XXY的个体表现为Klinefelter综合征；而47,XXX的个体则表现为超雌综合征。

近年的研究表明，X染色体失活的机制是相当复杂的。Matlei等认为，在人类X染色体Xq11.2～q21.1区域可能存在1个失活中心，X染色体的失活是从这个失活中心开始的。

（二）Y染色质

正常男性间期细胞用喹吖因等荧光染料染色，在荧光显微镜下可观察到细胞核内有1个直径约0.3 μm的强荧光小体，这是Y染色体长臂远端的异染色质，称为Y染色质(Y chromatin)或Y小体，细胞中Y染色质的数目与Y染色体数目相同(图2－11B)。例如，核型为47,XXY的个体，其X染色质与Y染色质的数目均为1；核型为47,XYY个体的X染色质数为0，Y染色质数为2。

临床通过性染色质检查，可诊断某些性染色体病；在产前诊断中，通过羊水细胞的性染色质检查可鉴定胎儿的性别，对防止性连锁遗传病患儿的出生有一定的意义。

小　结

细胞是一切生物进行生命活动的基本结构和功能单位。细胞内的遗传物质以染色质或染色体的形式存在。染色质和染色体实质上是同一物质在有丝分裂周期的不同阶段，执行不同生理功能时呈现的2种不同形式。染色质的主要成分是DNA和组蛋白。染色质以核小体作为基本结构，经螺线管、超螺线管，最后压缩包装成染色单体。根据间期细胞核内染色质的形态特征和染色性能。折叠盘曲程度及功能，将染色质分为常染色质和异染色质2种类型。

高等生物细胞增殖方式包括有丝分裂和减数分裂。

有丝分裂是体细胞增殖的主要方式，主要特点是DNA复制一次，细胞分裂一次；1个细胞分裂形成2个遗传组成相同的子细胞。细胞从前一次分裂结束开始生长到下一次分裂结束为止所经历的过程称为细胞周期，包括间期和分裂期，间期又分为G_1期、S期和G_2期3个分期。

减数分裂是有性生殖个体配子发生过程中进行的特殊方式的有丝分裂。其特点是DNA复制1次，细胞连续分裂2次，一个母细胞形成4个子细胞，子细胞的染色体数目减少一半。前期Ⅰ的同源染色体联会与互换，增加了配子的遗传多样性。精子与卵子是高度特化的单倍体生殖细胞，其形成过程称为配子发生。

人类性别是由参与受精的精子中性染色体类型决定的，但X和Y染色体在人类性别

决定中的作用并不相等。Y 染色体上的性别决定基因 SRY(Yp11.2)在性别决定中起关键作用。

根据着丝粒的位置,将人类染色体分为中央着丝粒染色体、亚中着丝粒染色体和近端着丝粒染色体 3 种类型。核型是指一个有丝分裂中期的体细胞内的全部染色体按一定形式排列起来所构成的图形。显带染色体有统一的识别和描述标准,记述染色体中某一特定带时需写明 4 方面的内容,即:①染色体号,②臂的符号,③区的序号,④带的序号。高分辨显带染色体的命名遵照 ISCN 体制。

性染色质是间期细胞核中性染色体的异染色质所显示出来的一种特殊结构,人类性染色质也包括 X 染色质和 Y 染色质 2 种。对于 X 染色质的形成机制,男性和女性 X 染色体基因产物量的一致性,Lyon 假说作出了较好的解释。

(金　洁)

第三章　遗传的分子基础

大量的研究证明,DNA是决定生物性状的分子基础。DNA既可以自我复制,也可以通过转录和翻译,控制细胞内蛋白质的生物合成,从而控制生物的性状。DNA还可能发生突变,使生物的性状发生相应的改变。

第一节　基因的概念

一、基因概念的提出

当代遗传学的奠基人孟德尔(Mendel),从1856年开始,以豌豆为实验材料进行了长达8年的杂交试验,于1866年在奥地利自然科学学会年刊上发表了著名的《植物杂交试验》,论文指出生物的每一个性状都是通过遗传因子(hereditary factor)传递的,遗传因子是独立的遗传单位;提出了遗传因子的分离和自由组合规律,证实了双亲传递给后代的不是性状而是遗传因子,遗传因子在世代传递中,其行为和功能具有相对的独立性。1903年,萨顿(Sutton)和鲍维里(Boveri)注意到在杂交试验中遗传因子的行为与减数分裂中染色体的行为非常吻合,提出了"遗传因子位于染色体上"的"萨顿－鲍维里假想"。1909年,丹麦生物学家约翰逊(Johannsen)将遗传因子更名为"基因(gene)"。自那时起,基因一词伴随着遗传学的发展沿用至今。

二、基因结构和功能的探索

摩尔根(Morgan)等人利用果蝇作了大量的潜心研究,于1926年出版了他的巨著《基因论》,建立了著名的基因学说,首次提出基因在染色体上按线形顺序排列,基因决定生物的性状,而且能发生突变并随着染色体同源节段的互换而交换。创立了遗传的染色体理论(chromosomal theory of inheritance)。1941年,比德尔(Beadle)等人提出了"一个基因一个酶"学说,证明基因通过控制酶的合成决定生化反应,进而决定生物的性状。1944年,艾弗里(Avery)等人首次证实DNA是遗传信息的载体。1953年,沃森(Watson)和克里克(Crick)根据X线衍射分析,提出了著名的DNA双螺旋结构模型,进一步说明基因成分就是DNA,它控制着蛋白质的合成。

1957年,本滋尔(Benzer)提出了顺反子学说,打破了关于基因是突变、重组、决定遗传性状的"三位一体"概念及基因是最小的不可分割的遗传单位的观点,认为基因是DNA分子上一段核苷酸序列,负责遗传信息的传递,一个基因可划分为顺反子、突变子和重组子等若干功能单位。1961年,雅各布(Jacob)等提出的乳糖操纵子学说,根据基因功能,把基因分为结构基因、调节基因和操纵基因等类型。1976年,美国科学家夏谱(Sharp)和英国科学家罗伯茨(Roberts)在对腺病毒的研究中,发现腺病毒结构基因中的编码序列被一些

非编码的 DNA 序列分隔开，提出了断裂基因的概念。1977 年，桑格(Sanger)等根据 ΦX174 噬菌体 DNA 碱基顺序揭示了基因的重叠性。

三、基因的现代概念

20 世纪 70 年代后，随着多学科的相互渗透和实验手段的更新，人们对基因的本质又有了更进一步的认识。在现代，人们对基因的认识，用分子生物学术语表达，即：基因是“合成有功能的蛋白质多肽链或 RNA 所需的全部核苷酸序列(通常是 DNA 序列)”。一个基因不仅包括编码蛋白质多肽链或 RNA 的核酸序列，而且包括为保证转录所必需的调控序列、5′端非翻译序列、内含子及 3′端非翻译序列等所有核苷酸序列。

基因的分类应该依据基因的功能来进行，由于大多数基因的功能尚不明了，因此目前还难于根据基因的功能分类。习惯上将基因分为以下几类：结构基因、调控基因、管家基因、奢侈基因及假基因等。

第二节　人类基因组

人类基因组是指人的所有遗传信息的总和，包括 2 个相对独立而相互关联的基因组(genome)：核基因组和线粒体基因组(参见第五章　线粒体遗传病)。如果不特别注明，通常所说的人类基因组是指核基因组。核基因组是指人类体细胞中 24 条不同染色体(22 对常染色体、X 和 Y 染色体)上的全部遗传信息，即：24 个 DNA 分子上 3.2×10^9 个核苷酸的遗传信息总和。人类基因组计划(human genome project，HGP)的初步研究结果表明：人类基因组中编码蛋白质和酶的结构基因为 3.5 万 ~ 4 万个，占人类基因组的 2% ~ 3%，95%以上为非编码序列。

根据 DNA 复性速度，人类基因组 DNA 显示出 3 个不同的时相，即分为 3 类：第 1 相复性速度最快，为高度重复序列；第 2 相复性速度较慢，为中度重复序列；第 3 相复性速度最慢，为单一序列。在人类基组中，单一序列约占 70%，重复序列约占 30%。

重复序列(repetitive sequence)是基因组中的多拷贝 DNA 序列，通常认为重复序列至少具有 20 个拷贝，多者可达几百个、上千个，甚至几十万个拷贝。重复序列中，少数属结构基因；另有一些与维持染色体结构的稳定性有关，如染色体端粒就是一种首尾衔接的(TAAGGG)*n* 重复序列；大多数重复序列的生物学功能尚不清楚，有待进一步研究。

一、单一序列

单一序列并非指单拷贝序列，实际上，单一序列包括惟一序列和低度重复序列。绝大多数编码各种不同功能蛋白质或酶的结构基因(structural gene)都是不重复的惟一序列，又称单拷贝序列(unique sequence)。单拷贝序列的长度通常在 1 000 bp 左右，故在全部 DNA 中所占的比例少。在单拷贝序列的两侧往往为散在分布的重复序列。由于绝大多数单拷贝序列编码蛋白质或酶，与人体各方面的生理功能和新陈代谢等密切相关。因此，这些序列的研究对医学实践有特别重要的意义。正是由于大多数结构基因的惟一性，导致了结构基因的突变很容易造成遗传性状的改变，或产生遗传病。

低度重复序列的拷贝数通常少于10个。因此,不同拷贝的基因产物在氨基酸的组成上存在差异。例如,人γ珠蛋白基因就是低度重复序列,包括$^{A}\gamma$和$^{G}\gamma$ 2个拷贝,二者的基因产物存在1个氨基酸的差异。

二、中度重复序列

中度重复序列(moderately repetitive sequence)多散在于人类基因组中,序列的长度和拷贝数很不一致,长度为300~7 000 bp,拷贝数少者为几十个,多者可达10^5个。大多数中度重复序列无编码功能,作为单拷贝基因的间隔序列,如*Alu*家族、*Kpn* I家族及假基因等。少数有编码功能,如多基因家族和超基因家族等。

(一)*Alu*家族和*Kpn* I家族

*Alu*家族又称*Alu*序列,是人类基因组中分布最广泛的一种中度重复序列,人类单倍体基因组中拷贝数达30万~50万,占人基因组的3%~10%。人类*Alu*序列长约300 bp,限制性核酸内切酶*Alu* I(AG↓CT)可将其切成130 bp和170 bp 2段,因而将其定名为*Alu*序列。*Alu*序列散在分布于人类基因组中,在间隔DNA和内含子中都发现有*Alu*序列,平均每5 kb DNA就有1个*Alu*序列。序列分析表明,*Alu*序列由两端的长约130 bp的正向重复序列及中间31 bp的插入序列三部分构成;*Alu*序列的5′端比较保守,但富含脱氧腺苷酸的3′端在不同的*Alu*序列变化很大;不同*Alu*序列的31 bp插入序列差异不大。*Alu*序列两侧为7~21 bp的正向重复序列,此正向重复序列在不同*Alu*序列各不相同。

*Alu*序列的功能现在还不清楚,原来有人推测其在hnRNA转录和加工中起作用,但现在的证据并不支持这一说法。另外,*Alu*序列中有一14 bp区域与一些感染哺乳动物的DNA病毒复制起始序列同源,因此推测*Alu*序列与DNA的复制也有关系,这还有待于进一步证实。

Kpn I家族与*Alu*家族不同,其平均长度为3 000~5 000 bp,拷贝数为3 000~4 000;因其含有限制性核酸内切酶Kpn I的切点而得名。*Kpn* I家族功能不明。

(二)多基因家族

多基因家族(multigene family)是中度重复序列中具有编码功能的基因家族,是指由一个祖先基因(ancestral gene)经过进化所形成的来源相同、结构相似、功能相关的一组基因,多基因家族是真核生物基因组最显著的特征之一。

按照基因的终产物,多基因家族分2类:一类编码RNA,包括*snRNA*基因、*tRNA*基因、*rRNA*基因等;另一类编码蛋白质,包括组蛋白基因、干扰素基因、珠蛋白基因及生长激素基因等。

按照在基因组中的分布,多基因家族也分为2类:一类是碱基序列基本相同的多个基因串联排列于同一条染色体上,形成基因簇(gene cluster),也称串联重复基因(tandemly repeat gene);另一类是多基因家族的不同成员成簇地分布在染色体的不同部位或者不同的染色体上,这些成员的序列虽然不完全相同,但是编码的蛋白质在功能上密切相关。

*rRNA*基因、*tRNA*基因及组蛋白基因等属于前一类。人类*rRNA*基因包括45 S *rRNA*基因和5 S *rRNA*基因。单倍体基因组中约含200个45 S *rRNA*基因,分布在13、14、15、21和22号5条近端着丝粒染色体的核仁组织区上,紧密排列,平均每条染色体约含40个拷

贝;5 S *rRNA* 基因大约有 1 000 个,全部位于 1q42 - 43;位于同一条染色体上的 1 000 ~ 2 000 个 *tRNA* 基因,为 50 ~ 60 种 tRNA 编码。在组蛋白基因家族,基因成簇地集中在 7q32 - q36 区域。基因组中存在大量重复序列用以编码组蛋白有重要意义,在 DNA 复制时,相应地组蛋白也成倍增加,而且往往在 DNA 合成一小段后,组蛋白马上与其结合,这要求在较短的时间内合成大量的组蛋白,因而需要有大量的组蛋白基因存在。

干扰素基因、珠蛋白基因、生长激素基因等属于后一类,珠蛋白基因分布于 α 珠蛋白基因簇和 β 珠蛋白基因簇,它们分别位于 16p13.11 - p13.33 和 11p15.5,这 2 个基因簇的每一个基因都含有 2 个内含子,虽然其长度不等,但位置非常相似。这 2 个基因簇的不同珠蛋白基因在序列编排上的高度一致性强烈地提示,它们是由一个祖先珠蛋白基因经过重复而来的。

在多基因家族成员中,有些成员并不产生有功能的基因产物,这些基因称为假基因或拟基因(pseudogene)。大多数多基因家族都有假基因存在,但在同一家族中,假基因数量只占少部分。假基因用 ψ 表示,例如 α 珠蛋白基因族 $\psi\alpha_1$ 表示与 α_1 相似的假基因。假基因与同源的有功能的基因结构十分相似,也有外显子和内含子,并且常与有功能的基因连锁在一起。但它们或者不能转录,或者转录后生成无功能的基因产物。这说明它们原来是有功能的基因,但在进化过程中,基因的关键部分发生了突变,如缺失、倒位或点突变等,这些突变引起转录调控失灵,或者阅读框架改变,或者氨基酸的密码子变成了终止密码子等,最终形成了假基因。

少数假基因缺少内含子,据推测,这种无内含子的假基因,可能是基因转录成的 RNA 前体,通过剪接失去内含子形成 mRNA,mRNA 经反转录产生 cDNA,再整合到基因组所致。例如,与 α 珠蛋白基因族 α_1 基因相比,假基因 $\psi\alpha_1$ 没有内含子。

(三)超基因家族

超基因家族(supergene family)是指一组由多基因家族及单基因组成的更大的基因家族。它们在结构上有程度不同的同源性,说明它们都起源于共同的祖先基因,但它们的功能并不相同,这是与多基因家族的差别所在。在一个超基因家族内,可含有几百个功能相关的基因。最经典的超基因家族是免疫球蛋白超基因家族,其成员都具有免疫球蛋白样的结构域。值得注意的是,应用现代生物信息学技术分析基因的核苷酸序列,已使越来越多的基因归为一类,从而有可能使原来多基因家族成为超基因家族。

三、高度重复序列

高度重复序列(highly repetitive sequence)是真核生物基因组与原核生物基因组的最大区别之一。人类高度重复序列的拷贝数在 10^6 以上,长度通常为 2 ~ 200 bp,在单倍体基因组中占 10% ~ 30%。高度重复序列中比较清楚的有 2 类:反向重复序列和卫星 DNA。

(一)反向重复序列

反向重复序列(inverted repeat sequence)复性速度极快,即使在极低的 DNA 浓度下,也能很快复性,因此倒位重复序列又被称为零时复性部分。反向重复序列由 2 个相同序列的互补拷贝在同一 DNA 链上反向排列而成,分为 2 种形式:一种是互补拷贝间有长约 1.6 kb 的序列间隔,这种结构在变性后再复性时,互补拷贝形成链内碱基配对,间隔部分膨

出，形成发夹式结构；另一种是互补拷贝间无间隔序列，称为回文序列（palindrome），回文结构长约 300 bp。反向重复序列约占人类基因组的 5%，大部分以单拷贝形式散布在基因组中，相邻反向重复序列间的平均间隔约为 12 kb。

（二）卫星 DNA

卫星 DNA（satellite DNA）在人类基因组中占 5% ~ 6%，串联排列，因在 CsCl 密度梯度离心时，在 DNA 主峰旁形成卫星状分布的小峰而得名。卫星 DNA 在人类基因组中分布广泛且表现出多态性，主要包括 2 类：6 ~ 25 bp 长的卫星 DNA 称小卫星 DNA（minisatellite DNA），2 ~ 6 bp 长的卫星 DNA 称微卫星 DNA（microsatellite DNA），又称为短串联重复（short tandem repeat，STR）。此外，在人类基因组中还存在有一类特殊的 α 卫星 DNA（α satellite DNA），分布在着丝粒区，最小重复单位为 171 bp。卫星 DNA 构成着丝粒、端粒、Y 染色体长臂上异染色质区以及随体区域。

第三节　基因的结构与功能

一、基因的结构

基因产物是蛋白质或酶的基因通常称为结构基因。真核生物结构基因与多数原核生物基因不同，原核生物的基因只有 1 条环状双链 DNA，编码序列是连续的，而真核生物结构基因的编码序列是不连续的。编码序列被非编码序列隔开，形成镶嵌排列的形式，因此结构基因又称为"断裂基因"（split gene）或"割裂基因"。值得注意的是，人类基因组中，有少数结构基因无内含子序列，如 *SRY* 基因和干扰素基因等；而在少数原核生物基因中也存在内含子。

（一）外显子和内含子

结构基因中的编码序列称为外显子（exon，E），相邻 2 个外显子之间的非编码序列称为内含子（intron，I）。不同结构基因所含外显子和内含子数目不同。例如，人 β 珠蛋白基因含 3 个 E 和 2 个 I，全长 1.7 kb，编码 146 个氨基酸（图 3－1）；人凝血因子基因Ⅷ含 26 个 E 和 25 个 I，全长 186 kb，编码 2 552 个氨基酸；而人假性肥大性肌营养不良症基因则含有 75 个外显子和 74 个内含子，全长 2 300 kb，编码 3 685 个氨基酸。

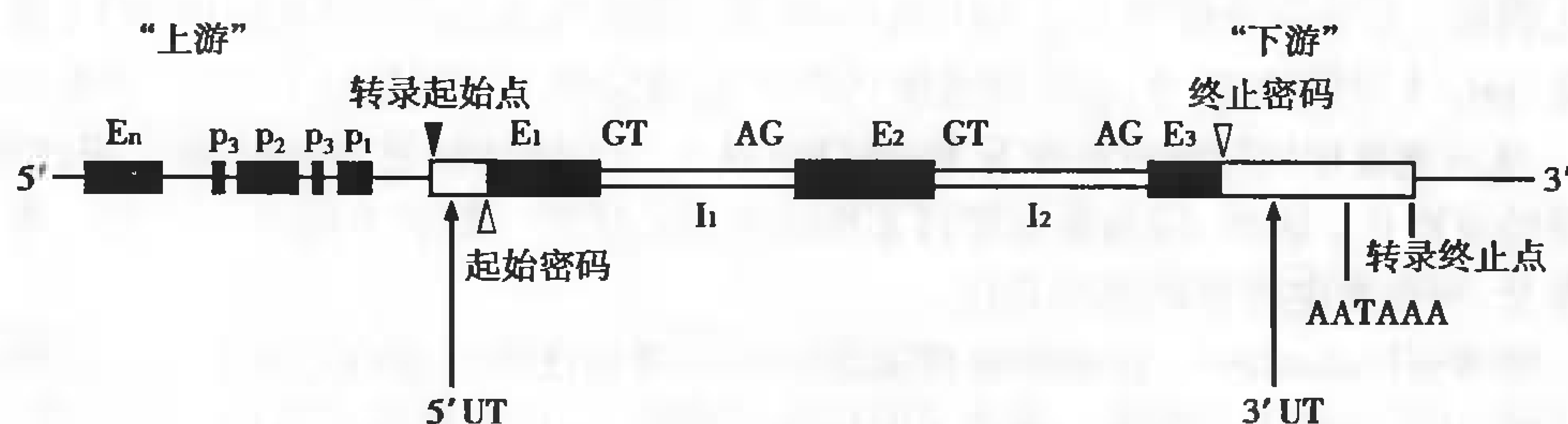


图 3－1　人 β 珠蛋白基因结构示意图

（二）外显子与内含子接头

所有结构基因的内含子5′端都以GT开始，3′端都以AG结束，即外显子和内含子接头区具有高度保守性，这种接头形式称为“GT－AG”法则，它是hnRNA加工时剪切信号。

结构基因中的内含子和外显子的关系不是固定不变的，有些DNA分子上的一段DNA序列在作为编码一条多肽链的基因时是外显子，而作为编码另一多肽链的基因时则是内含子，结果造成同一段DNA序列，可以转录出2条或2条以上不同的mRNA，使真核生物基因结构更复杂，同时也使有限的基因组中包含更多的遗传信息，这也是真核生物基因结构及其表达的重要特点。

（三）前导区

前导区（leader region）是位于转录起始点至第1外显子起始密码之间的一段序列，也称为5′端非翻译区（5′UT）。此段序列能够被转录，但不被翻译，转录后此段序列起维持mRNA稳定性的作用。

（四）调控序列

调控序列通常位于结构基因的“上游”和“下游”，故又称为侧翼序列（flanking sequence），对基因的表达起重要的调控作用。这些结构包括“上游”的启动子、增强子及“下游”3′端非翻译区（3′UT）的终止子等。

1.启动子（promotor） 是位于转录起始点上游约－100 bp左右的一段特异序列。RNA聚合酶的转录因子与此序列结合，方能启动基因的转录，但这段序列本身不被转录。常见的启动子序列包括TATA框、CAAT框和GC框。

（1）TATA框（TATA box） 是位于基因转录起始点上游－25～－30 bp处的一段高度保守序列。该序列（TATAA/TAA/T）的7个碱基中只有2个碱基（A/T，A/T）是可变的，周围为富含GC的序列。基因转录时，TATA框先与转录因子TFⅡ结合，然后再与RNA聚合酶Ⅱ结合并形成转录复合物，从而准确识别转录起始点，启动基因的转录。

（2）CAAT框（CAAT box） 位于转录起始位点上游－70～－80 bp处，也是一段保守序列，由9个碱基（GGC/TCAATCA）组成，其中只有1个碱基（C/T）是可变的。基因转移实验证实，CAAT框的存在是定量高效转录所必需的。

（3）GC框（GC box） GC框位于CAAT框的两侧，能与转录因子SP1结合，促进转录。

2.终止子（terminator） 位于结构基因3′端非翻译区（3′UT）的下游，在转录中提供转录终止信息。它由特异序列5′－AATAAA－3′及一段反向重复序列组成，AATAAA是多聚腺苷酸（poly A）的附加信号，能指导核酸内切酶在该信号下游特定位点裂解RNA并加上polyA。反向重复序列转录后形成发卡式结构（图3－2），该结构阻碍RNA聚合酶的移动，从而使转录终止。因此，反向重复序列又称转录终止信号。终止子的终止作用不是DNA序列本身，而是发生在转录成的RNA。

3.增强子（enhancer） 是指能够增强基因的转录活性的一段特定序列。当这些序列不存在时，转录水平大大降低。增强子的位置不固定，可位于转录起始位点上游或下游，甚至可以插入内含子中。增强子的功能与其位置和序列方向无关，可以是5′→3′，也可以是3′→5′。研究表明，增强子具有组织特异性，这是因为不同细胞核有不同的特异因子与增强子结合，从而对不同组织、器官的基因表达有不同的调控作用。例如，人类胰岛素基

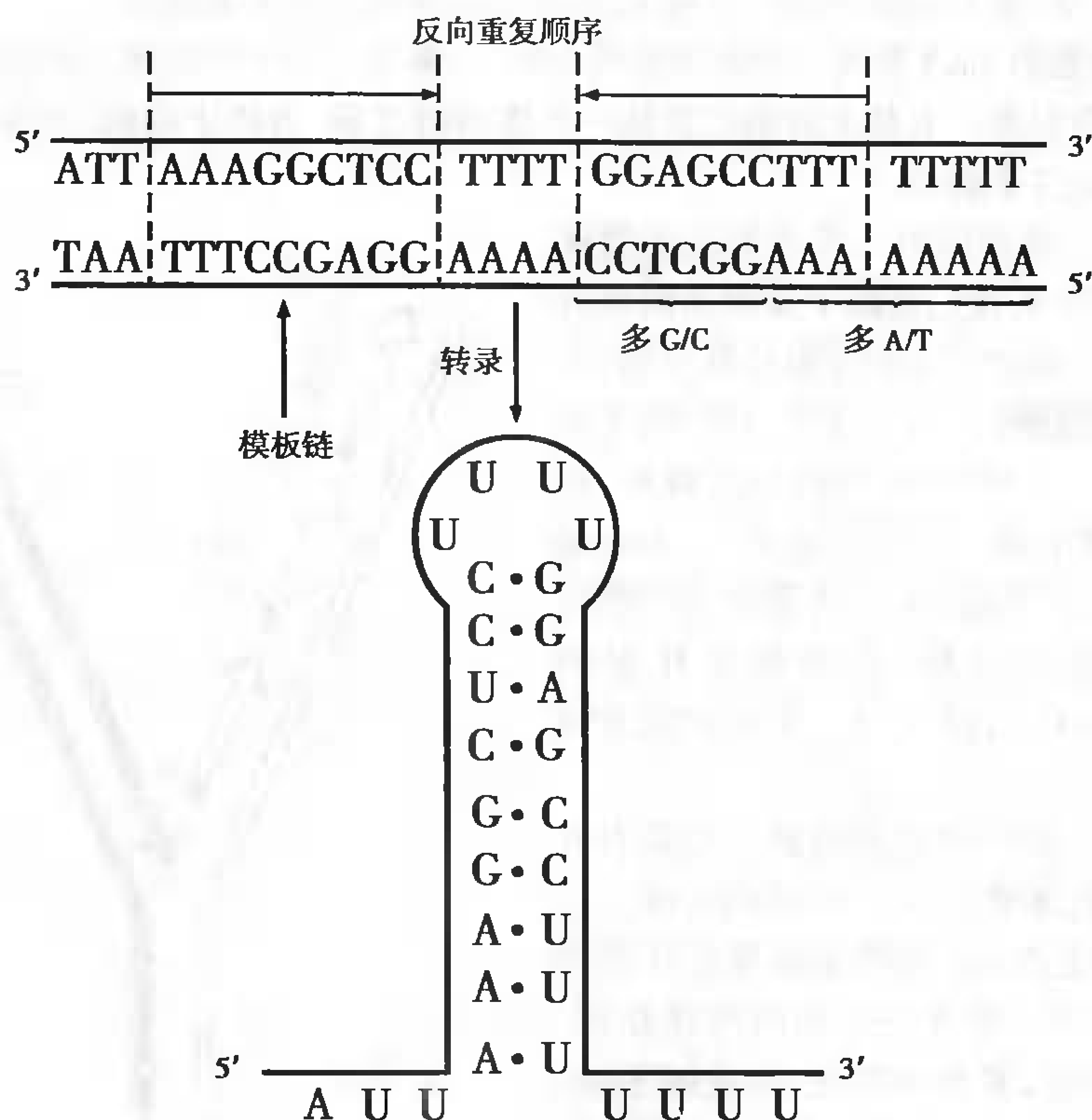


图 3－2 终止子的部分序列图解

因 5′末端上游约 250 bp 处有一组织特异性增强子，而在胰岛素 β 细胞中有一种特异性蛋白因子，该因子可增强胰岛素基因的转录。在其他组织细胞中没有这种蛋白因子，所以也就没有此作用。这就是为什么胰岛素基因只在胰岛 β 细胞中才高水平表达的重要原因。

二、基因的功能

基因的功能包括基因的复制和表达。基因复制的实质就是 DNA 复制（DNA replication）；基因表达（gene expression）是指 DNA 分子中蕴藏的遗传信息，通过转录传递给 RNA，再经翻译形成各种特异的蛋白质，从而使生物表现出千差万别的形态、生理特征和极为复杂的生命现象的过程。但 *rRNA* 基因和 *tRNA* 基因只转录不翻译。

（一）基因的复制

基因的复制是以 DNA 复制为基础的。遗传信息从 DNA 传递到 DNA，称为 DNA 复制。DNA 经过 1 个复制周期，形成的 2 个子代 DNA 分子，均由 1 条亲代的模板链和 1 条新链组成，因此，这种复制称为半保留复制（semiconservative replication）。

1. 复制的条件 DNA 复制需要以 4 种脱氧核苷三磷酸为原料（dATP、dGTP、dCTP、dTTP，总称 dNTP）；亲代的 DNA 双链作为模板；高能物质 ATP 供能及一系列酶作为催化剂。目前认为，原核细胞 DNA 复制需要 3 种 DNA 聚合酶（DNA polymerase，DNA－pol）和十

几种其他酶催化,真核细胞则需要5种DNA聚合酶和多种其他酶催化。

2.原核细胞的DNA复制 原核细胞的DNA复制是从一个特定起点同时向2个方向复制,称为双向复制。其整个复制过程是一个连续的过程,为便于理解,通常将其分为起始、延伸和终止3个阶段。

(1)起始 复制的第1步是辨认复制起始点。多种特定的蛋白质因子及酶准确辨认起始点序列,并与之结合构成起始复合物,继而Dna-B(解旋酶)、Dna-C(蛋白质)结合到起始点。Dna-B将DNA双链局部解开,从而在起点两侧形成"Y"形复制叉(replicating fork)(图3-3)。然后,在RNA聚合酶的作用下,以DNA链为模板,按碱基互补原则(A═U,G≡C),沿5′→3′方向合成RNA引物。

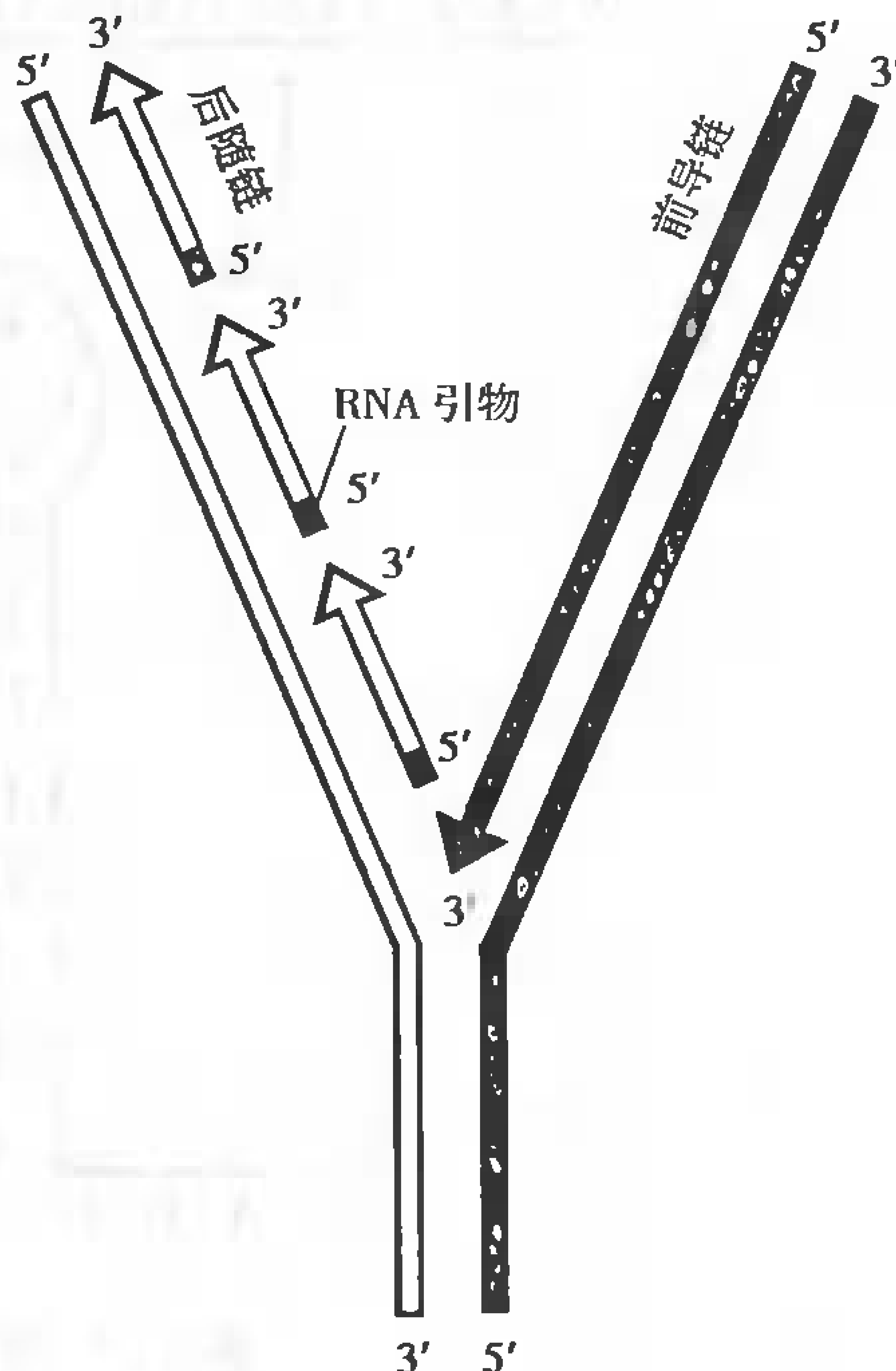


图3-3 DNA复制图解

(2)延伸 DNA聚合酶Ⅲ的作用具有方向性,只能催化新链沿5′→3″方向延伸。在3′→5′模板链上,DNA新链按碱基互补原则(A═T,G≡C)沿5′→3′方向连续复制,与解链方向一致,复制速度快,先完成复制,称为前导链(leading strand)。在5′→3′模板链上,DNA新链合成的方向与解链方向相反,且分段进行,每一段都需要一小段RNA引物;在引物引导下合成的DNA小片段称为冈崎片段(Okazaki fragment)。该条新链的复制不连续,速度慢,后完成复制,称其为后随链(lagging strand)。

(3)终止 延伸的DNA新链接近前方的RNA引物时,DNA聚合酶Ⅰ发挥5′→3′外切酶活性,RNA引物被水解。然后,DNA聚合酶Ⅰ发挥5′→3′聚合酶活性使新链继续延伸,填补引物水解后留下的空隙。最后,在DNA连接酶作用下,后一个冈崎片段的3′端与前一个岗崎片段的5′端或前导链的5′端连接起来,最终完成原核细胞的DNA复制。

3.真核细胞DNA复制的特点 真核细胞的DNA复制与原核细胞相似,如半保留复制、半不连续复制、双向复制等,但真核细胞DNA分子远比原核细胞大,而且DNA通常与组蛋白结合形成核小体,以染色质的形式存在于细胞核中。因此,真核细胞DNA复制与原核细胞有着明显的差异。

(1)多起点复制 真核细胞的每个DNA分子可有100~1 000个复制起点,包含1个复制起点、能够独立进行复制的复制单位称为复制子(replicon)。每个复制子只有起点,没有终点,随着复制的延伸,相邻的复制子汇合相连,最终完成复制。人类基因组约有10 000个复制子。原核细胞只有1个复制起点,即只有1个复制子。

(2)DNA链延伸速度慢 这是与真核细胞具有特殊的核小体结构不易解链有关。

(3)RNA 引物、冈崎片段小　真核细胞 DNA 复制的 RNA 引物约为 10 bp,原核细胞则可达数十 bp;在真核细胞,冈崎片段为 100 ~ 200 bp,而原核细胞则为 1 000 ~ 2 000 bp。

(4)DNA 聚合酶不同　真核细胞 DNA 复制起主要作用的 DNA 聚合酶为 DNA 聚合酶 α 及 DNA 聚合酶 δ,前者催化后随链的合成,后者催化前导链的合成。

(5)复制完成前不能开始下一轮复制　真核细胞 DNA 复制全部完成前,各复制起点不能开始下一轮的 DNA 复制,而快速生长的原核细胞则可在起始点上连续开始新的 DNA 复制。

(二)转录

遗传信息从 DNA 传递到 RNA,称为转录(transcription)。转录是以 DNA 分子中一条链为模板,在 RNA 聚合酶作用下,按照碱基互补原则,互补合成 RNA 分子的过程。

1.转录的条件　RNA 转录需要以 4 种三磷酸核苷(ATP、GTP、CTP、UTP,总称 NTP)为原料;以 3′→5′DNA 单链为模板,ATP 供能,RNA 聚合酶为催化剂。原核细胞只有 1 种 RNA 聚合酶;真核细胞有 3 种 RNA 聚合酶:RNA 聚合酶Ⅰ、RNA 聚合酶Ⅱ及 RNA 聚合酶Ⅲ。

DNA 双链中作为转录模板的单链称为模板链(template strand)或反义链(antisense strand),与其互补的另一条链则称为编码链(coding strand)或有义链(sense strand)(图 3 - 4A)。1 个双链 DNA 分子上有许多基因,并非每个基因的编码链都在同一条单链上。也就是说,DNA 双链某一段的编码链,在另一段则可以是模板链(图 3 - 4B)。

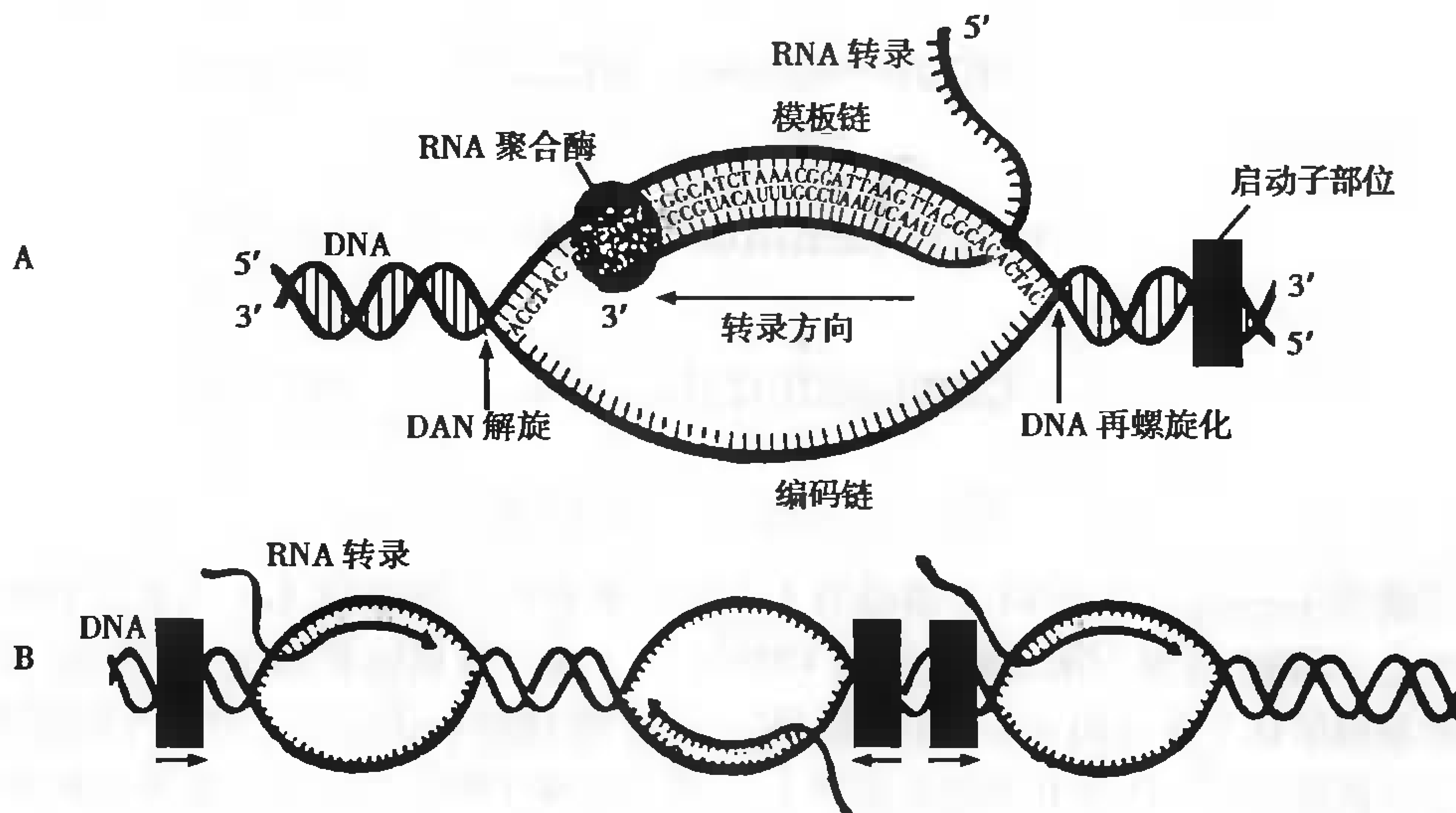


图 3 - 4　细菌 mRNA 的转录

2.原核细胞 RNA 的转录　催化原核细胞 RNA 转录的 RNA 聚合酶全酶由 5 条多肽链组成的核心酶加 σ 因子构成。①转录开始时,RNA 聚合酶全酶中的 σ 因子识别 DNA 分子上的启动子并与之结合,然后沿 DNA 链移动,使 DNA 双链局部解开,形成 12 ~ 17 bp 的单链区,此称为转录泡。②自转录泡处起,随着 DNA 聚合酶的移动,边解链边以 3′→5′DNA

链为模板，按碱基互补原则沿5′→3′方向转录RNA。RNA合成开始后，σ因子脱离RNA聚合酶。③随着转录向前推进，RNA链不断延长，完成转录的DNA区域又恢复双螺旋结构。④原核细胞转录的终止有2种机制，即依赖ρ因子（一种转录相关蛋白）的转录终止及不依赖ρ因子的转录终止。在前一种机制下，当DNA分子上出现转录终止信号时，ρ因子出现，它与核心酶结合使转录停止；ρ因子、核心酶及新合成的RNA单链与DNA链分离。

由于原核细胞没有核膜，且合成的RNA不需进行复杂的加工过程，故原核细胞转录和翻译2个过程是同时进行的。

3.真核细胞RNA的转录与加工 真核细胞RNA的转录是在细胞核中进行的，转录过程较为复杂，转录产物为各类RNA分子的前体，不具有活性，必须在核内经过复杂的加工修饰过程才能成为成熟的RNA分子，并从核孔输出到细胞质，参与蛋白质的合成。

(1)mRNA前体的加工 在真核细胞，结构基因的转录依赖RNA聚合酶Ⅱ的催化，转录的产物称为mRNA前体，又称为核内不均一RNA（heterogeneous nuclear RNA，hnRNA），其分子量是细胞质中成熟mRNA的7～10倍，且不具功能活性。需要经过戴帽、加尾、剪接等加工过程才能成为成熟的mRNA（图3－5）。

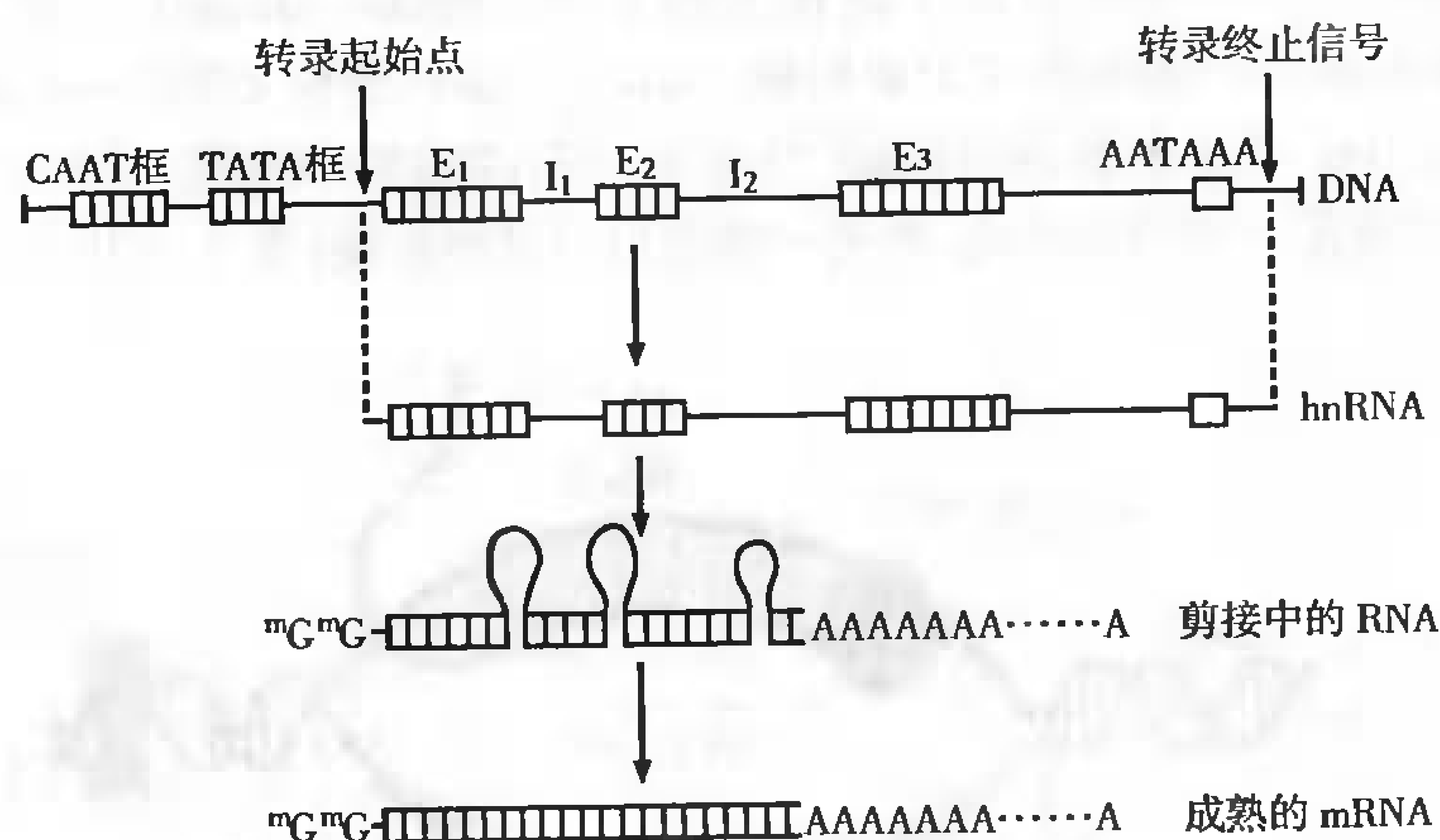


图3－5 mRNA前体的转录和加工

①戴帽（capping） hnRNA 5′端的第1个核苷酸通常为鸟苷酸（G），先由磷酸酶除去hnRNA的5′磷酸，在其5′加二磷酸并与GTP的α－P原子形成特有的5′－5′连接，再在鸟苷酸转移酶催化下使G的7′－N甲基化（5′－ m7G）形成帽0；再经2′－甲基转移酶催化在第1个核苷酸的2′－O上甲基化形成帽1；在第2个核苷酸的2′－O上甲基化形成帽2。“帽”的作用是使mRNA进入细胞质后，易被核糖体的小亚基识别，并与之结合。此外，帽还可起到封闭RNA的5′末端，使其免受磷酸酶和核酸酶的消化，增加其稳定性的作用。

②加尾（tailing） 是在hnRNA 3′端AAAUAA序列下游15～20 bp处，经加尾酶的作用，水解掉10～15个核苷酸，然后接上100～200个腺苷酸，形成多聚腺苷酸（poly A）“尾巴”的过程。poly A的功能主要是保持mRNA 3′端稳定，不受酶的破坏，并可促使mRNA由细胞核转运到细胞质中。

③剪接(splicing)　是指在酶的作用下,按“GT - AG”法则将 hnRNA 中的内含子切掉,然后把各个外显子按顺序连接起来的过程。

经上述 3 个环节加工形成的成熟 mRNA 由核孔进入细胞质,作为合成蛋白质的模板。

(2)tRNA 前体的加工　*tRNA* 基因的转录依赖 RNA 聚合酶Ⅲ的催化,转录产物 tRNA 前体的加工过程包括:3′端 CCA 取代 UU 构成作为氨基酸结合部位的氨基酸臂,切除 5′端的 16 个核苷酸先导序列(leader sequence),剪除内含子后将外显子连接,部分碱基被修饰为稀有碱基,使 tRNA 分子单链内部自身回折最终形成特殊的三叶草构型。

(3)rRNA 前体的加工　真核细胞的 *rRNA* 基因包括 45 S *rRNA* 基因和 5 S *rRNA* 基因。45 S *rRNA* 基因在 RNA 聚合酶Ⅰ的催化下,转录合成 45 S rRNA 前体,然后经一系列加工过程,最后生成 18 S、5.8 S 和 28 S 3 种 rRNA。5 S *rRNA* 基因在 RNA 聚合酶Ⅲ的催化下,转录合成 5 S rRNA。5 S rRNA、28 S rRNA 和 5.8 S rRNA 参与核糖体大亚基的组装,而 18 S rRNA则参与小亚基的形成。

(三)翻译

翻译(translation)是指以 mRNA 为模板合成蛋白质多肽链的过程。

1.遗传密码　DNA 分子中所蕴藏的遗传信息,通过碱基互补配对原则转录传递给 mRNA。DNA 或 mRNA 分子的 4 种碱基随机组成 64 个密码子,其中 61 个为编码氨基酸的密码子,AUG 为起始密码(在以哺乳动物为代表的真核生物中 AUG 编码蛋氨酸,在以细菌为代表的原核生物中 AUG 则编码甲酰蛋氨酸);有 3 个密码子不编码任何氨基酸,它们是蛋白质合成的终止信号,即终止密码 :UAA、UAG、UGA。

遗传密码具有如下特点。

(1)通用性　自然界所有生命形式,都共用相同的一套遗传密码。但近年来的研究发现,动物细胞的线粒体和植物细胞的叶绿体中的遗传密码与“通用密码”不完全相同。

(2)简并性　在 64 个密码子中,61 个为编码氨基酸的密码子,而组成蛋白质的氨基酸只有 20 种,这说明存在有 2 个或多个密码子编码同一种氨基酸的简并现象。

(3)方向性和连续性　翻译总是从 mRNA 分子 5′端 AUG 开始,按 5′→3′方向连续阅读,无间隔区。

(4)摆动性　编码氨基酸的密码子共 61 个,而 tRNA 的数目却少于 61 种,这说明有的 tRNA 能识别不同的密码子,即密码子的第 3 位碱基与反密码子的第 1 位碱基配对不严格,此称为摆动配对。

(5)偏倚性　虽然一个氨基酸可以有多个密码子,但对一个特定的氨基酸而言,有的密码子使用频率明显高于其他密码子,这就是所谓的遗传密码使用的偏倚性。

2.氨基酸的激活　合成蛋白质所需要的氨基酸必须是活化的氨基酸,因为氨基酸本身不能识别遗传密码,而必须与相应的 tRNA 结合形成氨酰 - tRNA 复合体,通过 tRNA 的反密码子为媒介辨认 mRNA 上的密码子。氨基酸在氨酰 - tRNA 合成酶的催化下与 ATP 反应,生成氨酰 - AMP,然后,氨酰 - AMP 把氨基酰基团转移到 tRNA 分子上,形成氨酰 - tRNA复合体,此过程称为氨基酸的激活。氨酰 - tRNA 合成酶具有高度的专一性,即一种酶只能活化一种氨基酸。

3.多肽链的合成　多肽链的合成是在 mRNA、tRNA 和核糖体三者密切配合下完成

的。mRNA 是合成多肽链的模板或信息载体，tRNA 能转运氨基酸，核糖体提供多肽链合成的场所。现以原核细胞为例介绍蛋白质的合成过程。

(1)起始　翻译开始时，在起始因子(initiation factor，IF)的作用下，核糖体的 30 S 小亚基结合到 mRNA 分子的起始密码 AUG 部位，形成“30 S－mRNA－IF”复合体。然后，甲酰蛋氨酰－tRNA 进入复合体，tRNA 分子上的反密码子与 mRNA 分子上的起始密码 AUG 互补结合。最后，核糖体的 50 S 大亚基与小亚基结合形成 70 S 起始复合体，起始因子释放，翻译开始阶段完成。

(2)延长　核糖体上有 2 个氨酰－tRNA 的结合位点：供体部位(P 位)和受体部位(A 位)。在 70 S 起始复合体中，甲酰蛋氨酰－tRNA 结合在 P 位，第 2 个氨酰－tRNA 结合在 A 位。在转肽酶催化下，P 位上的氨酰－tRNA 所携带的氨酰基移至 A 位，与 A 位上的氨基酰脱水缩合形成二肽。空载的 tRNA 离开 P 位。在移位酶和 GTP 供能的情况下，核糖体沿 mRNA 5′→3′方向移动 1 个密码子，即 A 位上的氨酰－tRNA 移位到 P 位。空载的 A 位又可接受 1 个新的氨酰－tRNA，不断重复此过程，肽链不断增长。所合成的多肽链中，氨基酸的种类、数目及排列顺序完全由 mRNA 上密码子决定。

(3)终止　当 mRNA 分子上出现终止密码(UAA、UAG、UGA)时，核糖体停止移动，肽链停止延长。在释放因子的参与下，多肽链从核糖体上释放。同时，核糖体的大、小亚基分离，mRNA 与核糖体分离。

蛋白质合成速度很快，大肠杆菌合成 1 个 300～400 个氨基酸的蛋白质仅需要 10～20 s。通常 1 个 mRNA 分子可以同时结合多个核糖体，即同一 mRNA 可依次翻译出多条相同的多肽链。

三、基因表达的调控

在原核生物，转录和翻译是同步进行的；但在真核细胞，结构基因的转录和转录后的加工是在细胞核内进行，而翻译则是在细胞质内进行。即基因表达是在不同时间、不同地点进行的。

理论上，改变遗传信息传递过程的任何环节均会影响基因的表达。转录过程的许多环节均可作为调控点，是基因表达调控最重要、最复杂的一个层次。在真核细胞，转录后加工过程也是基因表达的重要调节方式。在翻译过程中，影响蛋白质合成的因素，同样也是调节基因表达的因素。翻译后的加工可直接、快速地改变蛋白质的结构和功能，是细胞对外环境变化或者某些特异刺激应答时的快速反应机制。总之，在遗传信息传递过程中的各个环节，各个水平均存在基因表达的调控机制。

基因表达调控的研究是目前基因研究的热点。

(一)原核细胞基因表达的调控

1961 年，Jacob 和 Monod 首先在研究大肠杆菌乳糖代谢过程中发现了基因表达的调控系统，提出了大肠杆菌乳糖操纵子模型，开创了基因表达调控研究的新纪元。现以大肠杆菌乳糖代谢的调控为例，介绍原核细胞基因表达的调控过程。

1.操纵子和乳糖操纵子　操纵子(operon)是指功能上相关的一组基因串联组成的一个转录单位，由调控元件和结构基因两部分组成。调控元件位于转录起始点附近，包括与

RNA 聚合酶结合的启动子、与阻遏蛋白(repressor)结合的操纵基因(operator,O)。结构基因是相互串联的1至几个基因。因此,操纵子既是基因的转录单位,又是基因表达调控的功能单位。

乳糖操纵子的结构基因由 *Lac Z*、*Lac Y* 和 *Lac A* 3个彼此相连的基因组成，分别编码大肠杆菌利用乳糖的3种酶:β-半乳糖苷酶、透性酶及乙酰基转移酶。调控元件的操纵基因、启动子位于结构基因的上游,操纵基因长约38 bp,含有1个反向重复序列,有26 bp与启动子(825 bp)相重叠。启动子 TTGACA 序列和 TATAATG 序列是 RNA 聚合酶识别和结合部位,后者类似于真核细胞的 TATA 框。在操纵基因上游较远处还存在调节基因(regulatory gene,R),编码阻遏蛋白,该蛋白与操纵基因结合则可阻止 RNA 聚合酶与启动子结合,关闭启动子,阻碍结构基因的表达。

2.乳糖代谢的调控 当在大肠杆菌的培养基中加入乳糖时,乳糖分子进入大肠杆菌细胞内并与阻遏蛋白结合,引起后者的构象变化并离开操纵基因。原来处于关闭状态的乳糖操纵子变为开放状态,RNA 聚合酶与启动子结合,开始转录和翻译,产生3种乳糖代谢所需的酶,这一过程称为诱导(图3-6)。合成的3种酶促使乳糖分解成葡萄糖和半乳糖,被细菌吸收和利用。在诱导过程中,乳糖为诱导物,乳糖加入细菌培养基中2~3 min后,3种酶分子便形成。

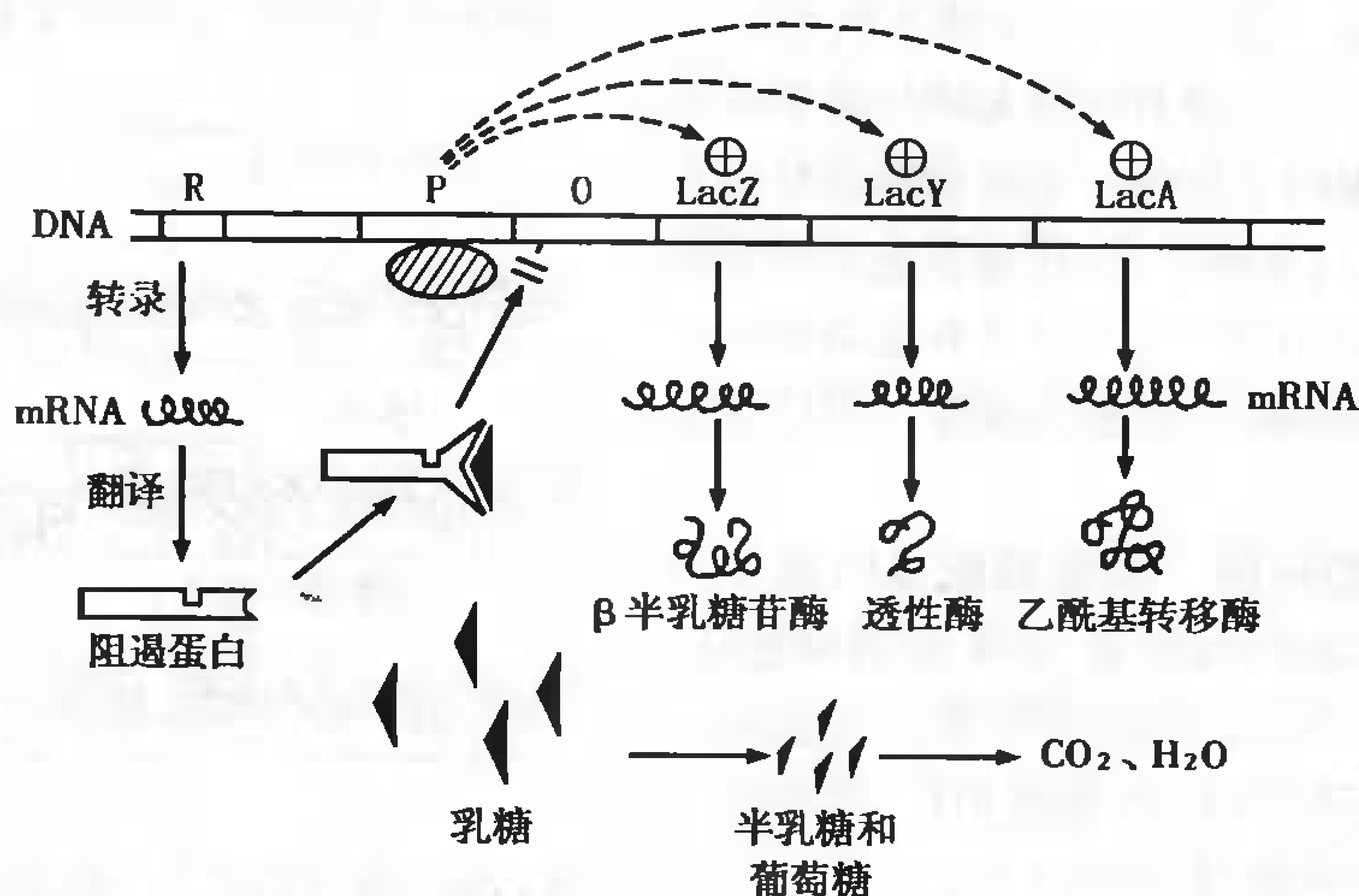


图3-6 乳糖操纵子的激活(诱导)

当乳糖分解完毕,阻遏物恢复原来的构象而与操纵基因结合,使乳糖操纵子处于关闭状态,3个结构基因停止转录,此过程称为阻遏(图3-7)。

大肠杆菌乳糖操纵子的基因表达调控机制表明,原核生物在环境诱导物存在的条件下,基因的相互作用使代谢系统中的酶系统能够按照所需要的数量精确合成;没有环境诱导物存在,相应的基因关闭,不产生相应的酶,使其能更好地适应环境。这是原核生物进化过程中所获得的调节功能,体现了其对环境的最佳适应能力。真核生物尽管也有适应环境的要求,但基因调控最重要的特点、最直接的作用则是调节个体生长发育,使遗传信息在特定时间、特定位置规律性地表达。

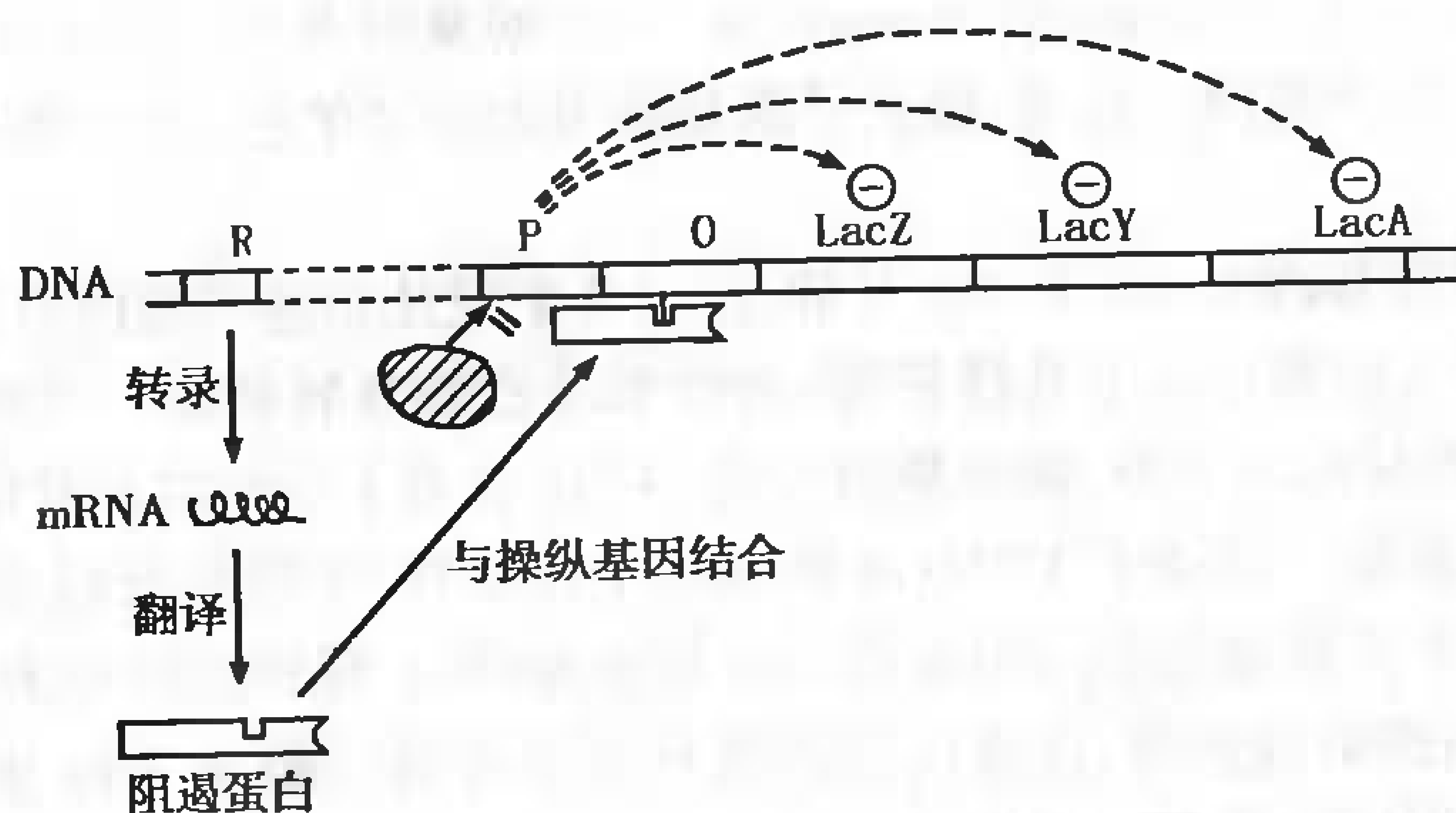


图3-7 乳糖操纵子的关闭(阻遏)

(二)真核细胞基因表达调控

真核细胞基因含有大量非编码序列,转录后形成的RNA需经加工才能成为成熟的mRNA。绝大多数真核生物行有性生殖,体细胞含2个基因组。生物体的不同组织器官的细胞虽然含有相同的基因组,但在个体发育的不同阶段,细胞内的基因表达是不同的。如胎儿蛋白基因仅在胚胎期表达,出生后不表达;珠蛋白基因,有些在胚胎前期表达,有些则在胚胎后期表达。此外,同一个体不同组织器官的细胞中基因表达的种类和数量也存在明显的差异。例如,大鼠肝细胞基因转录的量是肾细胞和脾细胞的2倍多。真核细胞遗传信息量大,合成的蛋白质种类多,其基因表达的调控在多层次、多水平上进行,并且具有复杂的时空性。因此,真核细胞基因表达调控比原核生物复杂得多。

1.转录前的调控 真核细胞基因组DNA与组蛋白、非组蛋白及少量RNA组成染色质,且染色质有一定程度的盘绕折叠。组蛋白与DNA结合,既可保护DNA免受损伤,维持基因组的稳定性,也可抑制基因的表达。

染色质重组实验证明,将染色质中的组蛋白用酸除去后,残余的染色质成分表现出更高的转录活性;当组蛋白成分回复后,染色质的转录活性降低。这表明组蛋白能抑制DNA的转录活性。非组蛋白在间期核中与DNA结合能解除组蛋白对DNA转录的抑制,促进转录,这表明非组蛋白是重要的基因调控成分。Stein和Kleinsmith于1975年提出“组蛋白转位模型”,用以说明真核细胞的基因调控(图3-8)。该模型

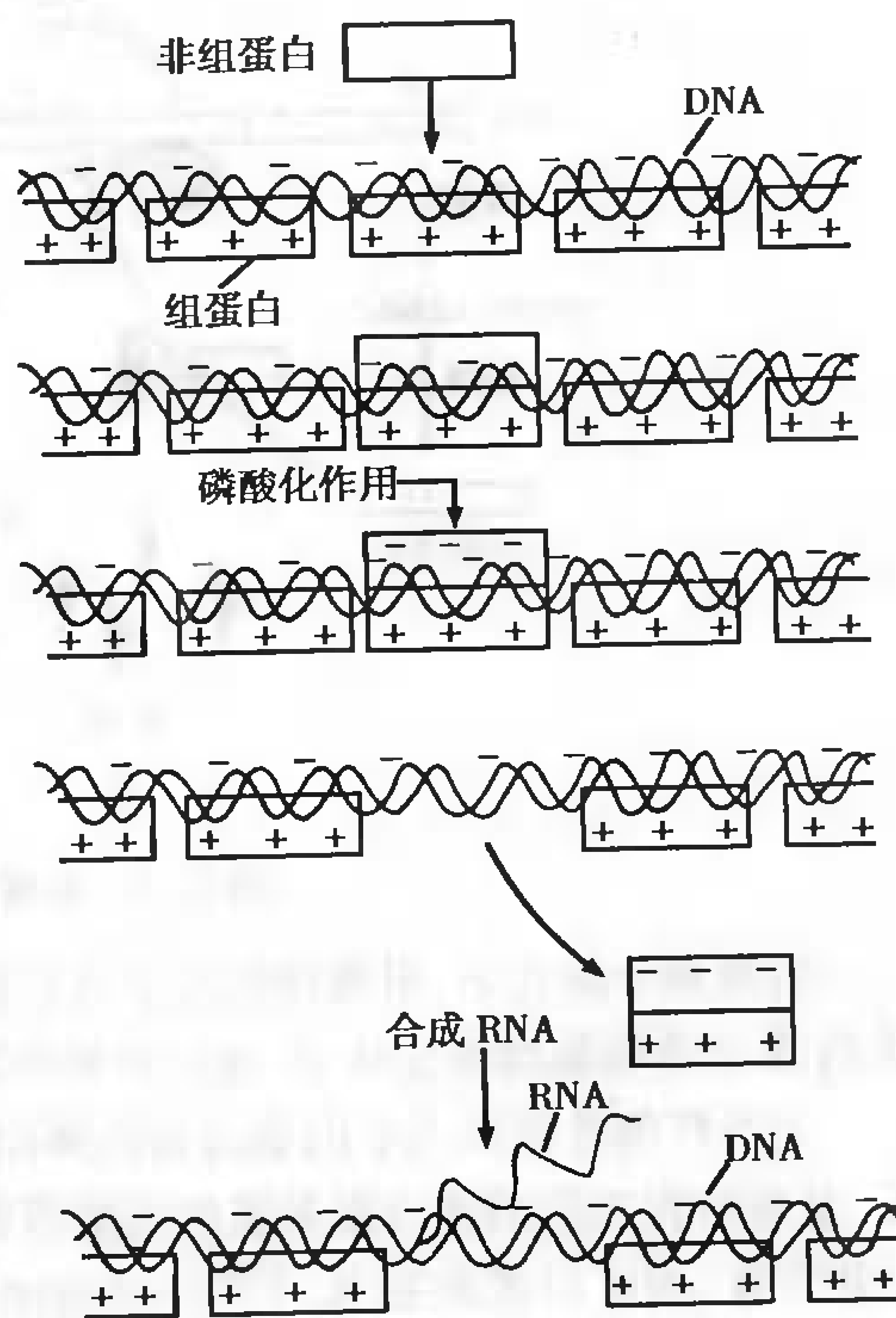


图3-8 组蛋白转位模型

认为,构成染色质的结构单位核小体是由带正电荷的组蛋白与带负电荷的 DNA 结合而成,组蛋白与 DNA 结合抑制了基因的转录。一些特异的非组蛋白能识别 DNA 的特异片段并与之结合,由于磷酸基带负电荷,当某种非组蛋白磷酸化后,带负电荷的磷酸化非组蛋白与带正电荷的组蛋白结合成复合物,并与组蛋白一起从 DNA 上脱离,解除了组蛋白对 DNA 的抑制,DNA 便开始转录。如果复合物中非组蛋白去磷酸化,组蛋白便与之分离,重新与 DNA 结合,使活化的 DNA 受到抑制,转录停止。非组蛋白的磷酸化还受细胞内 cAMP 浓度的影响,cAMP 通过激活蛋白激酶使非组蛋白磷酸化而参与调节基因的表达。

2.转录水平的调控　真核细胞基因转录调控是通过顺式作用元件和反式作用因子相互作用实现的。顺式作用元件(cis – acting elements)是指那些与结构基因表达调控相关的、能够被基因调控蛋白特异性识别并与之结合的 DNA 序列,包括启动子、增强子及反应元件等。反应元件存在启动子附近或增强子内,能介导基因对细胞外的某些信息分子产生反应。反式作用因子(trans – acting factor)包括与顺式作用元件特异结合的一些转录因子。只有当各类转录因子与启动子或增强子、转录因子之间以及转录因子与 RNA 聚合酶之间相互作用,才能启动基因的转录。转录因子的大多数是正性调节因子,起激活转录的作用;少数是负性调节因子,起抑制转录作用。转录因子又受激素、病毒、金属离子、光、热、射线及化学物质的影响而改变其作用,或改变转录因子与特异调控序列结合的能力。因此,转录水平的调控是复杂而关键的调节环节,受多种因素的影响。

3.转录后的调控　转录形成的 hnRNA 长度比成熟 mRNA 长得多,它既有编码序列,又包括内含子,需加工为成熟的 mRNA,并由细胞核转至细胞质,方能作为模板参与蛋白质的合成。加工过程中的选择性剪接(alternative splicing)的效率以及戴帽、加尾等过程都受到调控并决定转录后 mRNA 的特征。

4.翻译水平的调控　翻译水平的调控主要包括 mRNA 的稳定性、翻译的准确性、翻译效率等。这些过程受 mRNA 的成熟度、核糖体的数量、始动因子(IF)、延长因子(EF)、释放因子(RF)以及各种酶的影响。例如,IF 的活性是通过其磷酸化调节的,一些致癌剂如佛波酯(phorbol ester)或癌基因产物等,通过信号传递途径或作为这一途径成分,调节 IF 的磷酸化;而一些病毒(如脊髓灰质炎病毒)也可以使某些 IF 降解而失去活性。

5.翻译后的调控　翻译后形成的初级产物需要在细胞质中加工、修饰才能成为有活性的、成熟的蛋白质。有的初级产物是一条多肽链,经剪接加工、组装后形成由几条肽链构成的蛋白质,胰岛素的加工过程便是如此;有的初级产物加工后,又可形成多种具不同功能的蛋白质,如促黑色素激素和 β – 内啡肽等是由一条多肽链的初级产物经剪接加工后形成的;还有的初级产物则需要化学基团的修饰。

综上所述,真核细胞基因表达调控是多层次、多水平、多因素、网络性的调控过程,但很多调控环节、调控机制还不清楚,有待进一步研究。

第四节　基因突变

基因突变是指基因结构的碱基对组成或排列顺序的改变。基因突变是新基因产生的方式,基因突变后在原座位上出现的新基因,称为突变基因(mutation gene);而未发生突变

的基因称为野生型(wild type)基因。如果突变基因改变了原有蛋白质的结构或功能,将引起个体表型的变化。

一、基因突变的特性

(一)可逆性

基因突变的可逆性是指野生型基因可以突变形成突变型基因,突变型基因也可突变为野生型基因。前者称正向突变,后者称回复突变。当正向突变和回复突变的频率相等时,群体基因频率不发生变化。但正向突变和回复突变的频率一般不同,往往是正向突变频率高于回复突变频率,从而使群体基因频率发生变化。

(二)稀有性

基因突变普遍存在于自然界中,从病毒到人类都有基因突变发生。不同种生物及同种生物的不同基因突变率各不相同。自然情况下基因突变率很低,一般只有百万分之几,用 $n \times 10^{-6}$/(基因·代)表示。

(三)多向性

基因突变的多向性是指一个基因可以向多个不同的方向突变,形成复等位基因。复等位基因是指同一基因座位上有 2 个以上的基因存在。例如,人类 ABO 血型基因位于 9q34.1,该基因位点上 I^A、I^B 和 i 3 个基因构成一组复等位基因,决定 ABO 血型抗原的合成。又如,控制果蝇眼睛颜色的一组复等位基因多达 14 个。

(四)有害性和有利性

大多数的突变都是有害的,不利于生物的生长发育。因为任何一种生物的遗传基础都是经历了长期自然选择的结果,具有一定的适应性,突变打破了机体的这种协调平衡关系,对生物有机体产生有害效应。

但基因突变也存在有利的一面,因为有少数突变更有利于生物的生存。如农业生产上利用基因突变培育的优质高产品种,提高了农作物抗病、抗倒伏及抗病虫害的能力;在工业微生物育种中应用射线诱发青霉菌基因突变培育出了高产青霉素的优良菌株等。

二、基因突变的诱因

基因突变可以是在自然条件下发生的,也可是人为因素所引起。前者称为自发突变,后者称为诱发突变。

(一)自发突变

自发突变(spontaneous mutation)又称自然突变,产生的原因很多:可能是自然环境中的辐射本底和致突变剂引起 DNA 损伤,也可能是机体在正常代谢过程中产生的自由基以及糖基化终末产物引起 DNA 损伤。另外,DNA 复制过程中碱基错误配对(简称错配)也可引起基因突变。人体约有 6×10^{14} 个细胞,每个细胞含有 6×10^9 碱基对,人的一生约有 10^{16} 次细胞分裂,故发生错配的机会多。但细胞内存在一系列使 DNA 保持高度稳定性的机制。研究表明,大肠杆菌 DNA 复制过程的错配率为 $10^{-1} \sim 10^{-2}$,DNA 聚合酶 3′端或 5′端外切酶可切除错配碱基,使 DNA 复制错配率降到 $10^{-5} \sim 10^{-6}$。此外,复制后的校正系统

可使 DNA 复制错配率进一步降至 10^{-9}。人类的 DNA 复制过程导致的错配,经校正可使复制错配率降低至 10^{-10}。

(二)诱发突变

诱发突变(induced mutation)的因素有物理因素、化学因素及生物因素等。

1.物理因素 可以导致基因发生突变的物理因素包括电离辐射、紫外线及温度等。电离辐射带有较强的能量,作用于生物机体,可直接使受照射的物质发生电离产生自由基,后者与细胞内的 DNA 分子发生激烈的化学反应,引起 DNA 主链及染色体断裂。这种畸变一般不能被修复或发生错误的修复。紫外线的能量较低,对 DNA 的损伤主要是使链上相邻的嘧啶碱基之间通过共价结合形成嘧啶二聚体(T－T、C－C、C－T),也可引起 DNA 与 DNA 交联或 DNA 与蛋白质交联,甚至可致 DNA 链断裂。嘧啶二聚体使 DNA 双链间的氢键减弱,导致 DNA 结构局部变形,影响 DNA 的复制和转录。通常细胞内的 DNA 修复系统可修复这种 DNA 损伤,如果修复系统异常,不能修复受损的 DNA 分子,则产生基因突变。

2.化学因素 能引起基因突变的化学因素很多,如甲醛、氮芥、烷化剂、亚硝酸、亚硝酸盐及苯并芘等。食品中含有的亚硝酸盐进入人体后发生亚硝化反应生成亚硝胺,后者可使 DNA 分子中的嘧啶结构改变,从而使 DNA 在复制中发生配对错误,导致基因突变,引发消化器官癌变。烟草中的多环苯蒽在芳烃羟化酶作用下转变为多环芳烃,后者可致基因突变,引发肺癌、喉癌及口腔癌等。

3.生物因素 能诱发基因发生突变的生物因素主要是病毒,包括 DNA 病毒和 RNA 病毒。DNA 病毒可直接整合到宿主细胞的 DNA 分子中,引起基因突变或细胞癌变;RNA 病毒可在反转录酶的作用下以病毒 RNA 为模板,反向合成 cDNA,cDNA 可整合到宿主细胞的 DNA 分子中引起基因突变或癌变。

三、基因突变的机制

基因突变的方式有多种,人类基因组中常见的突变有碱基替换、移码突变及动态突变等。基因突变不仅可以发生在编码区,也可以发生于非编码区,包括启动子、终止子及剪接部位等。基因突变可以发生于核 DNA,也可发生于线粒体 DNA。

(一)碱基替换

DNA 分子中的一个核苷酸被另一个核苷酸所替代,称为碱基替换(base substitution),又称点突变(point mutation)。碱基替换包括转换和颠换 2 种方式。转换(transition)是指 DNA 分子中的一种嘌呤被另一种嘌呤或一种嘧啶被另一种嘧啶替换;颠换(transversion)是指一种嘧啶被另一种嘌呤或一种嘌呤被另一种嘧啶取代。自然界中的点突变,转换多于颠换。点突变发生后,必然导致相应 mRNA 分子密码子的改变。根据密码子改变引起的效应不同,点突变又分为以下几种。

1.同义突变(samesence mutation) 是指基因突变后的密码子所编码的氨基酸与原来密码子编码的氨基酸相同。因此,同义突变不产生突变效应。例如,DNA 中的密码子 GCG 的第 3 位的 G 被 A 取代形成 GCA,经 GCG 和 GCA 转录来的 mRNA 中的 CGC 和 CGU 均编码精氨酸(图 3－9)。

2.错义突变(missense mutation) 是指基因突变导致编码一种氨基酸的密码子变为编码另一种氨基酸的密码子,从而使所合成的蛋白质分子改变(图 3-9)。大多数错义突变对机体产生不利影响。

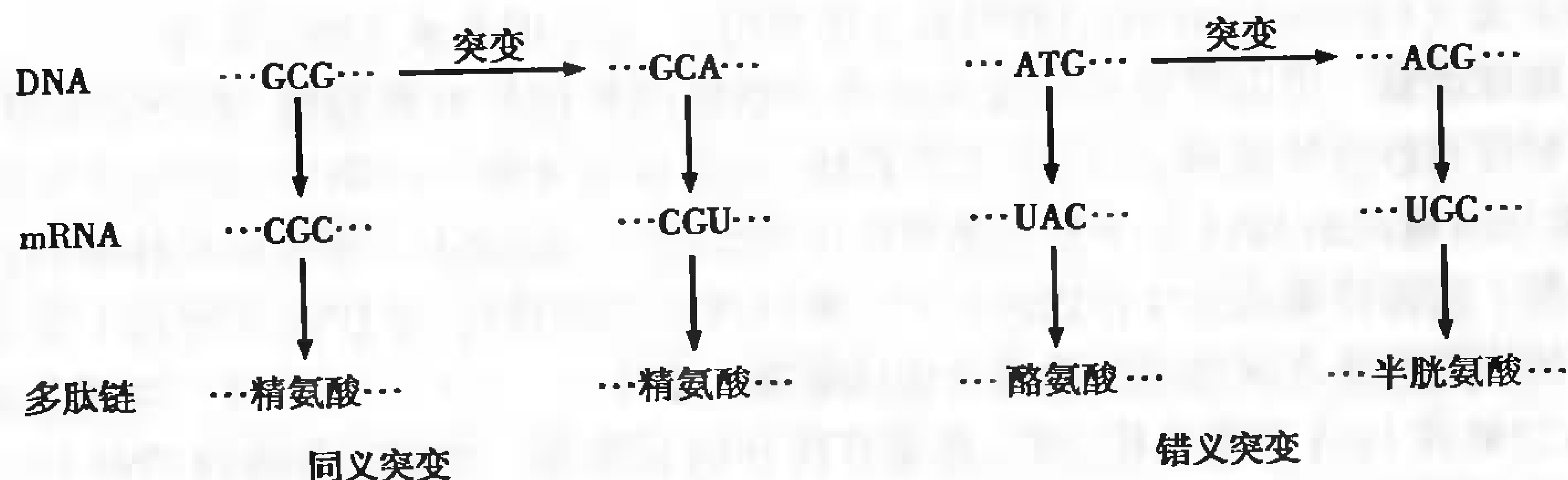


图 3-9 同义突变与错义突变

3.无义突变(nonsense mutation) 是指点突变使原来某一编码氨基酸的密码子变为终止密码,导致多肽链合成提前终止,肽链缩短,成为无活性的多肽片段(图 3-10)。例如β珠蛋白基因第 17 位密码子产生无义突变,导致β珠蛋白肽链缩短,功能异常,形成β^0珠蛋白生成障碍性贫血(β^0地中海贫血)。

4.终止密码突变(termination codon mutation) 是指点突变发生在基因的终止密码上,使终止密码变为编码某一氨基酸的密码子,导致多肽链合成继续进行,直至下一个终止密码出现才停止。这样形成的多肽链比正常多肽链长,故也称为延长突变(elongation mutation)。

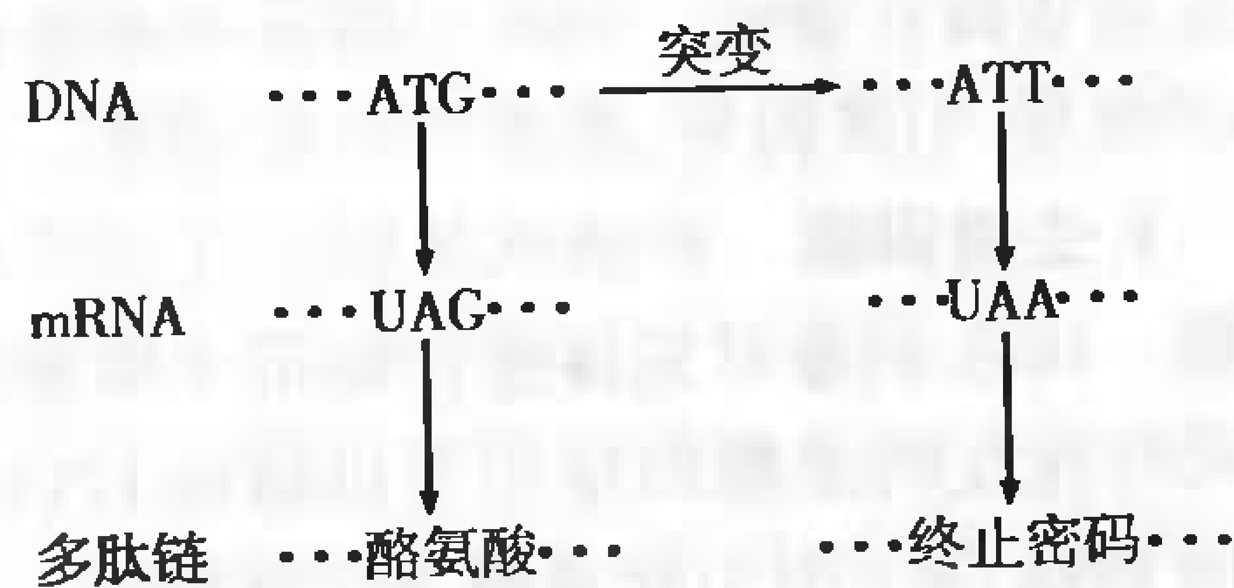


图 3-10 无义突变

(二)移码突变

移码突变(frame shift mutation)是指 DNA 分子中插入或缺失 1 个或几个碱基对,从而使这一位置下游的 DNA 读码顺序改变所造成的突变(图 3-11)。碱基的插入或缺失可以是 1 个或几个碱基对,也可以是十几个甚至几千个碱基对的大片段,这种突变将导致蛋白质的严重改变,引起严重的疾病。例如,假性肥大性肌营养不良症(DMD)基因可有几个 kb 的缺失。

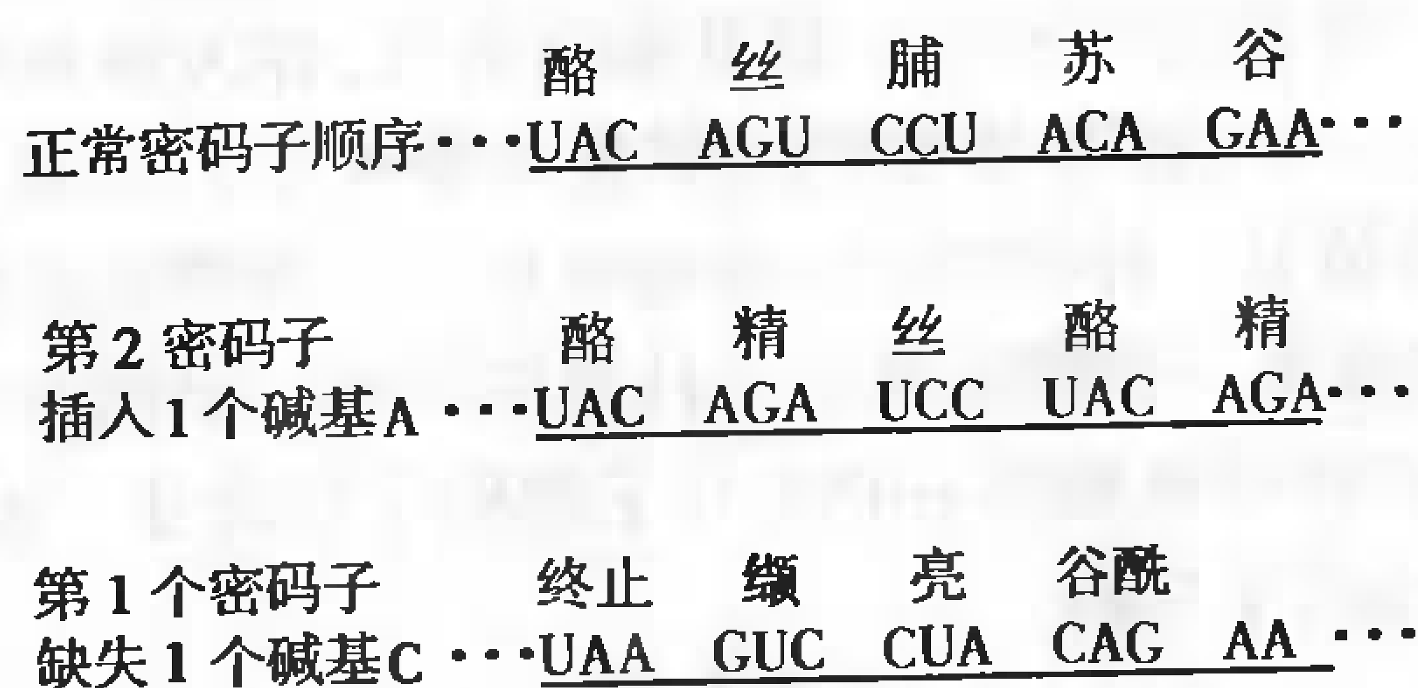


图 3-11 移码突变

(三)动态突变

动态突变(dynamic mutation)是指人类基因组中短串联重复序列(STR),尤其是三核苷酸重复序列,在靠近基因或位于基因序列之中时,其重复次数在一代一代的传递过程中不断增加的现象。动态突变是导致人类遗传病的一种新的基因突变类型,如脆性 X 综合征。位于 X 染色体上的脆性 X 综合征的致病基因 *FMR* - 1 的 5′端及其上游的非编码区有一段不稳定的(CGG)n 三核苷酸重复序列。正常人的(CGG)n 拷贝数为 6 ~ 54,当拷贝数增加至 54 ~ 200 时称为前突变,为无临床表现的携带者;当(CGG)n 拷贝数超过 200 称为全突变,出现智能低下及其他的脆性 X 综合征的特征,现已知重症脆性 X 综合征患者(CGG)n 拷贝数最多可达 2 000。目前已发现人类的 20 多种遗传病是由于动态突变所引起。动态突变揭示了导致人类遗传病新的发病机制,解释了遗传早现和不完全外显的分子机制,为相关疾病提供了简便有效的基因诊断方法。

四、基因突变的表型效应

基因的核苷酸序列决定蛋白质的结构。基因核苷酸序列的稳定性是蛋白质(或酶)稳定性的基础。基因结构的改变将产生不同程度的表型效应,根据基因突变对机体的影响程度,将基因突变的表型效应分为以下几类。

(一)对机体不产生有害效应

包括几种情况:①突变后的密码子与突变前的密码子编码同一氨基酸,没有引起该多肽链中氨基酸顺序和种类的改变;②不影响基因功能的重复序列和基因间隔序列的突变;③某些错义突变虽然导致蛋白质中氨基酸组成改变,但不影响机体的生理功能,只是形成正常人体生化组成的遗传学差异。例如,人类的血清蛋白类型、ABO 血型、HLA 类型及各种同工酶型等都是基因突变形成的。这种遗传学差异通常对人体没有明显的有害效应,但在某些情况下也会产生严重后果。例如,不同血型间输血、不同 HLA 型间的器官移植排斥反应等。

从进化观点看,上述类型的基因突变称为中性突变(neutral mutation)。这是生物多样性与进化发展的重要源泉,为生物进化提供了丰富的原料。在人类基因组中,这类基因突变是 DNA 多态性(DNA polymorphism)的遗传基础。

(二)产生有利于生物有机体生存的积极效应

这类基因突变能促进或加强某些生命活动,更有利于生物的生存。例如,人类基因突变产生的镰形细胞贫血症,纯合子(HbSHbS)患者病情重,杂合子 (HbAHbS) 患者大部分无明显症状。在非洲黑种人中,杂合子患者比正常人(HbAHbA)更具抗恶性疟疾的能力。因此,在某种意义上,杂合子患者的基因突变更有利于个体生存。

(三)产生遗传病

基因突变影响个体的生长发育,导致生育能力降低,寿命缩短。据估计,一个人至少带有 5 ~ 6 个处于杂合状态的有害突变,当这些突变处在纯合状态时就会产生有害后果。人类各种遗传病的基因最初都是由正常基因突变来的,编码结构蛋白的基因突变产生分子病,如血红蛋白病;编码酶蛋白的基因突变,使酶功能异常或缺如,导致代谢过程紊乱或

代谢中断，产生先天性代谢病，如苯丙酮尿症。另外，基因突变与肿瘤形成密切相关，某些基因产物参与其他基因的表达调控，调节细胞的生长、分化，这类基因突变导致其功能降低或丧失，则可能导致肿瘤发生。

基因突变可以发生在个体发育的任何时期，发生在不同发育时期的基因突变对个体的影响不同。发生在胚胎细胞的基因突变，将严重影响胎儿的生长发育，导致流产或死胎或胎儿畸形。发生于体细胞的基因突变，在行无性繁殖的生物中可通过细胞分裂传递给后代；行有性生殖的生物，体细胞基因突变不传递给后代，故不会引起后代遗传性状的改变，但可导致当代体细胞癌变。若基因突变发生于生殖细胞，将随生殖细胞（精子和卵）传递给后代，引起后代遗传性状改变或产生遗传病。

小　结

DNA 是控制生物性状遗传的分子基础，基因是 DNA 分子上具有特定遗传效应的核酸片段，它决定细胞内 RNA 和蛋白质（包括酶分子）的合成，从而决定生物的遗传性状。

人类基因组包括核基因组和线粒体基因组，每个核基因组 DNA 长约 3.2×10^9 bp。人类基因组中编码蛋白质和酶的结构基因数量为 3.5 万 ~ 4 万个，约占人类基因组的 2% ~ 3%，95%以上为非编码序列。在序列组成上，人类基因组的单一序列约占 70%、重复序列约占 30%；重复序列又分为中度重复序列和高度重复序列。

真核生物的结构基因是断裂基因，它是由编码序列和非编码序列镶嵌排列形成的断裂结构。断裂基因中的编码序列称为外显子，在外显子之间起间隔作用的非编码序列称为内含子。外显子和内含子之间存在有特殊的接头顺序“GT – AG”。每个断裂基因的首尾外显子的外侧都有一段调控序列或称侧翼序列（启动子、增强子和终止子），它们对基因的表达起着重要的调控作用。启动子位于结构基因转录起始点上游，启动子序列包括 TATA 框、CAAT 框和 GC 框。终止子位于结构基因的下游，由 5′ – AATAAA – 3′特定序列及一段反向重复序列组成，在转录中提供转录终止信息。增强子能够增强基因的转录活性。

DNA 分子可以自我复制，也可以经过转录将特定的遗传信息传递给 RNA，再经过翻译控制细胞内蛋白质的合成。

基因突变是指基因结构的碱基对组成或排列顺序的改变。按照诱因，将基因突变分为自发突变和诱发突变两类，诱发基因突变的因素包括物理因素、化学因素及生物因素等；基因突变具有可逆性、多向性、稀有性、有害性及有利性；基因突变的机制包括碱基替换、移码突变及动态突变等，其中碱基替换又包括同义突变、错义突变、无义突变和终止密码突变几种形式。基因突变对机体影响不同，大多数基因突变影响个体的生长发育，产生遗传性疾病；少数基因突变对机体不产生有害效应，有些基因突变更有利于机体的生存。

（梁素华）

第四章　单基因遗传病

单基因遗传病是指由一对等位基因异常所引起的疾病。单基因遗传病的遗传规律符合孟德尔定律，故又称为孟德尔式遗传病。根据致病基因所在染色体（常染色体或性染色体）和基因性质（显性或隐性）的不同，将单基因遗传病分为：常染色体显性遗传病、常染色体隐性遗传病、X连锁显性遗传病、X连锁隐性遗传病及Y连锁遗传病。线粒体遗传病也属于单基因遗传病，其致病基因位于线粒体内，呈现特殊的母体遗传。

第一节　单基因遗传病的遗传方式

对人类单基因遗传病或性状的研究常采用系谱分析法（pedigree analysis）。系谱是反映某种遗传病在一个家系中发病情况的分布图解，系谱中第1个被医生或研究人员发现的患者称为先证者（proband）。系谱分析常从先证者入手，追溯调查家系中所有的人员，包括直系亲属和旁系亲属的发病情况及婚配情况，再将调查资料绘制成系谱并取系谱进行分析。绘制系谱常用的符号见图4－1。

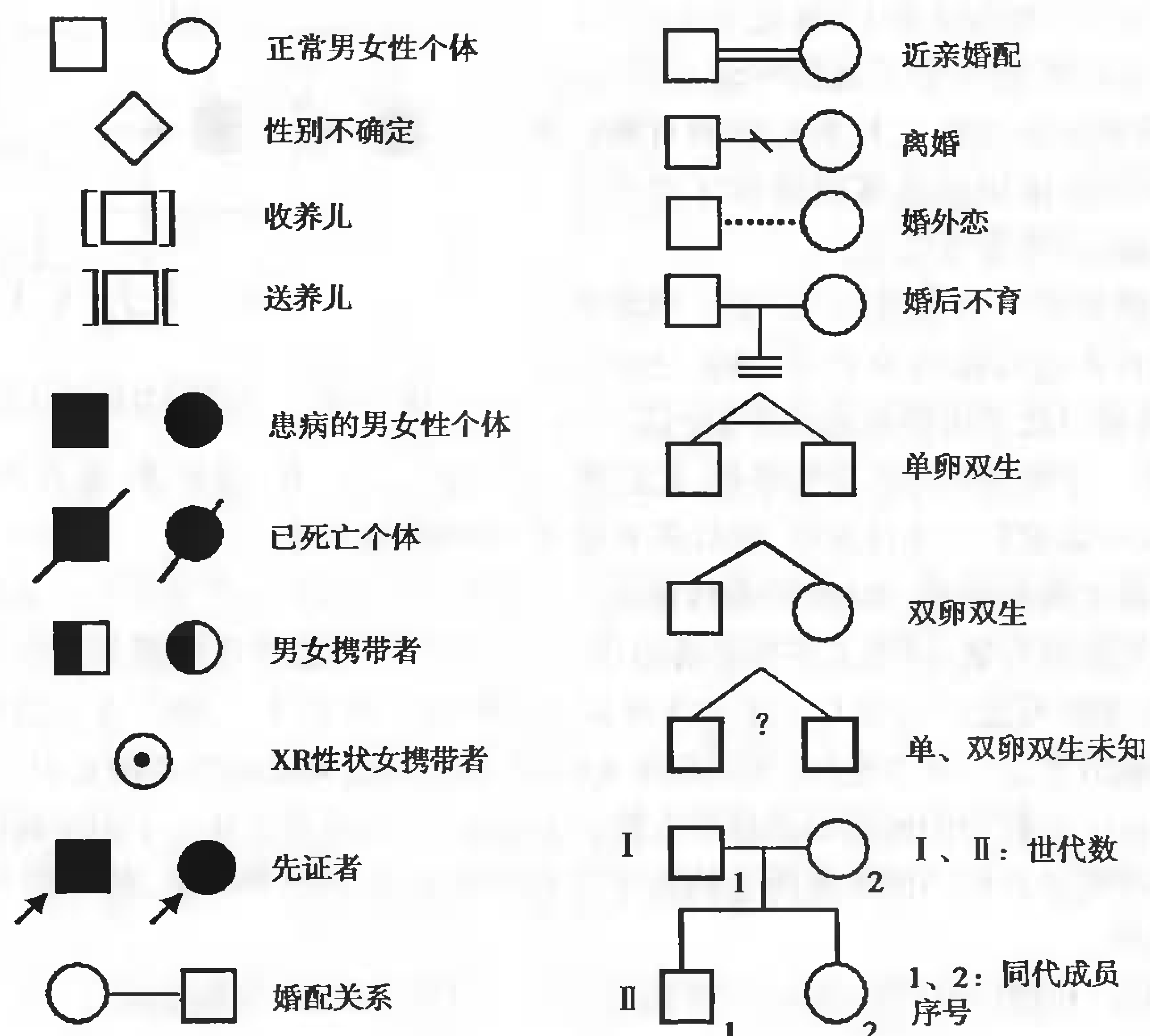


图4－1　系谱中常用符号

一、常染色体显性遗传

如果控制某种性状或疾病的基因位于常染色体(1～22 号染色体)上,而且基因是显性的,这种遗传方式就称为常染色体显性遗传(autosomal dominance inheritance, AD)。符合 AD 方式的疾病称为常染色体显性遗传病。在 AD 中,根据杂合子与显性纯合子的表型是否一致,将其分为以下几种类型。

(一)完全显性

在常染色体显性遗传中,如果杂合子的表型与显性纯合子完全一致,就称为完全显性(complete dominance)。家族性多发性结肠息肉(FPC)是一种常染色体显性遗传病,该病的致病基因位于 5q21。FPC 患者在青少年时期结肠和直肠上长有多发性息肉(良性肿瘤),但随年龄的增长逐渐恶变,最终形成结肠癌。由于患者常出现血性腹泻,故常被误诊为肠炎,90%未经治疗的患者死于结肠癌。

致病基因往往是由正常基因突变产生的,在 AD 病,正常基因为隐性基因(a),致病基因为显性基因(A)。由于基因突变率很低,一般为 10^{-5}～10^{-6}/(基因·代),故人群中致病基因的频率很低(0.01～0.001)。因此,在 AD 病中,显性纯合子(AA)患者罕见,绝大多数患者是杂合子(Aa)。患者(Aa)与正常人婚配(aa),每个子女都有 1/2 的可能性为患者。

图 4－2 是一例家族性结肠息肉的系谱,系谱中先证者Ⅱ$_3$ 的结肠息肉已经恶变为结肠癌,而且术后复发,他的母亲Ⅰ$_2$ 和姐姐Ⅱ$_1$ 均死于结肠癌。先证者的 3 个子女尚年幼,未出现临床表现,但他们的结肠上长有息肉的可能性为 1/2。故家系中未出现临床表现的年幼个体应该及早就诊,以防癌变发生。

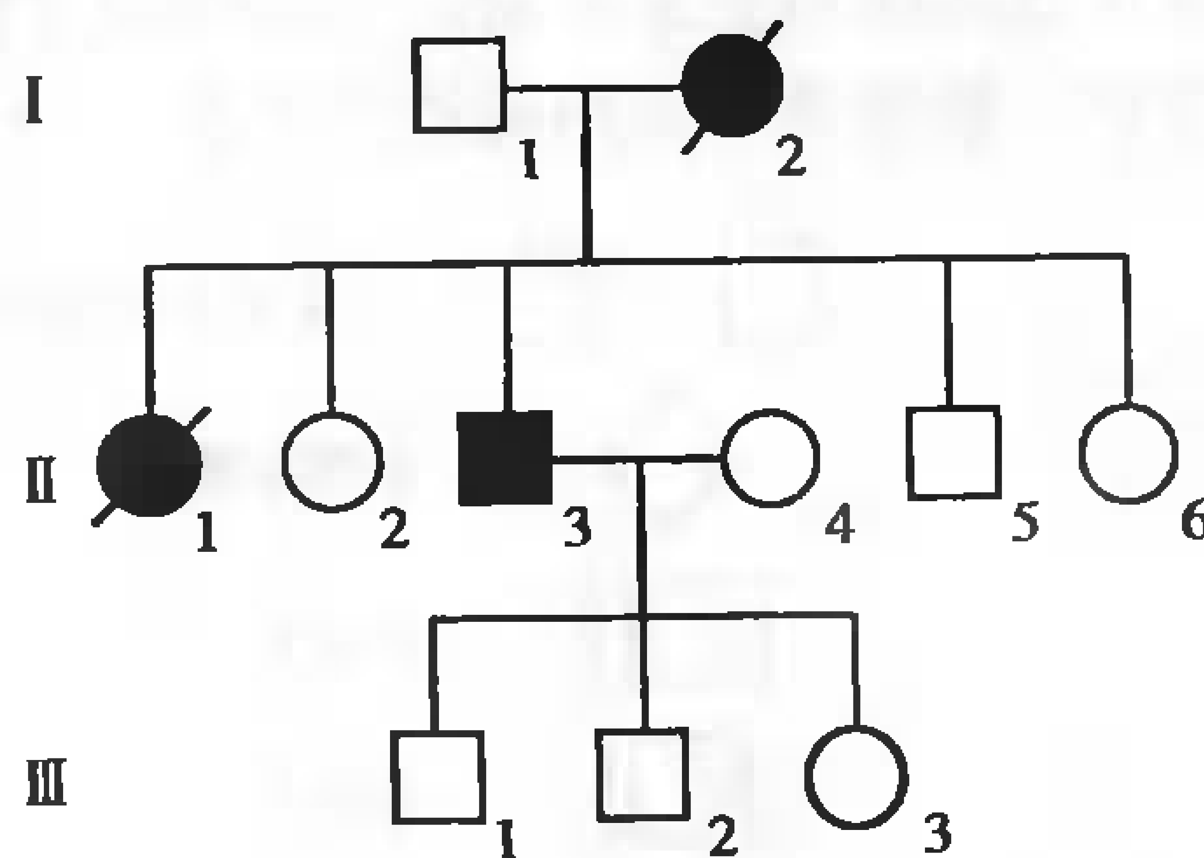


图 4－2　一例家族性结肠息肉系谱

牙本质发育不全也属于 AD 病。患者的牙齿上往往有灰色或蓝色的乳光出现,牙釉质早期脱落,容易引起牙齿磨损而致牙冠凹陷变短。图 4－3 是一牙本质发育不全的家系,先证者Ⅲ$_2$ 的双亲之一Ⅱ$_6$ 是患者,患者的后代中有患者,而且在该家系的 4 代人中,每代都有患者,发病率为 1/2。

根据以上典型系谱,AD 病系谱具有以下主要特点:①通常患者双亲之一是患者,由于致病基因的频率很低,故绝大多数患者为杂合子;②系谱中连续几代都有患者;③患者的同胞、子女患病可能性均为 1/2;④由于致病基因位于常染色体上,故男女发病机会均等;⑤双亲无病时子女一般不发病,如果双亲无病而子女发病,则是新发生的基因突变所致。

在生育子女数少的家庭中容易产生较大的偏差,很难看到上述 1/2 的发病比例;但如果将若干婚配方式相同的夫妻所生的孩子发病情况进行统计学分析,就可得到近似 1/2 的发病比例。

常见的 AD 病还有多指(趾)症、并指(趾)症、短指症及亨廷顿舞蹈病等。

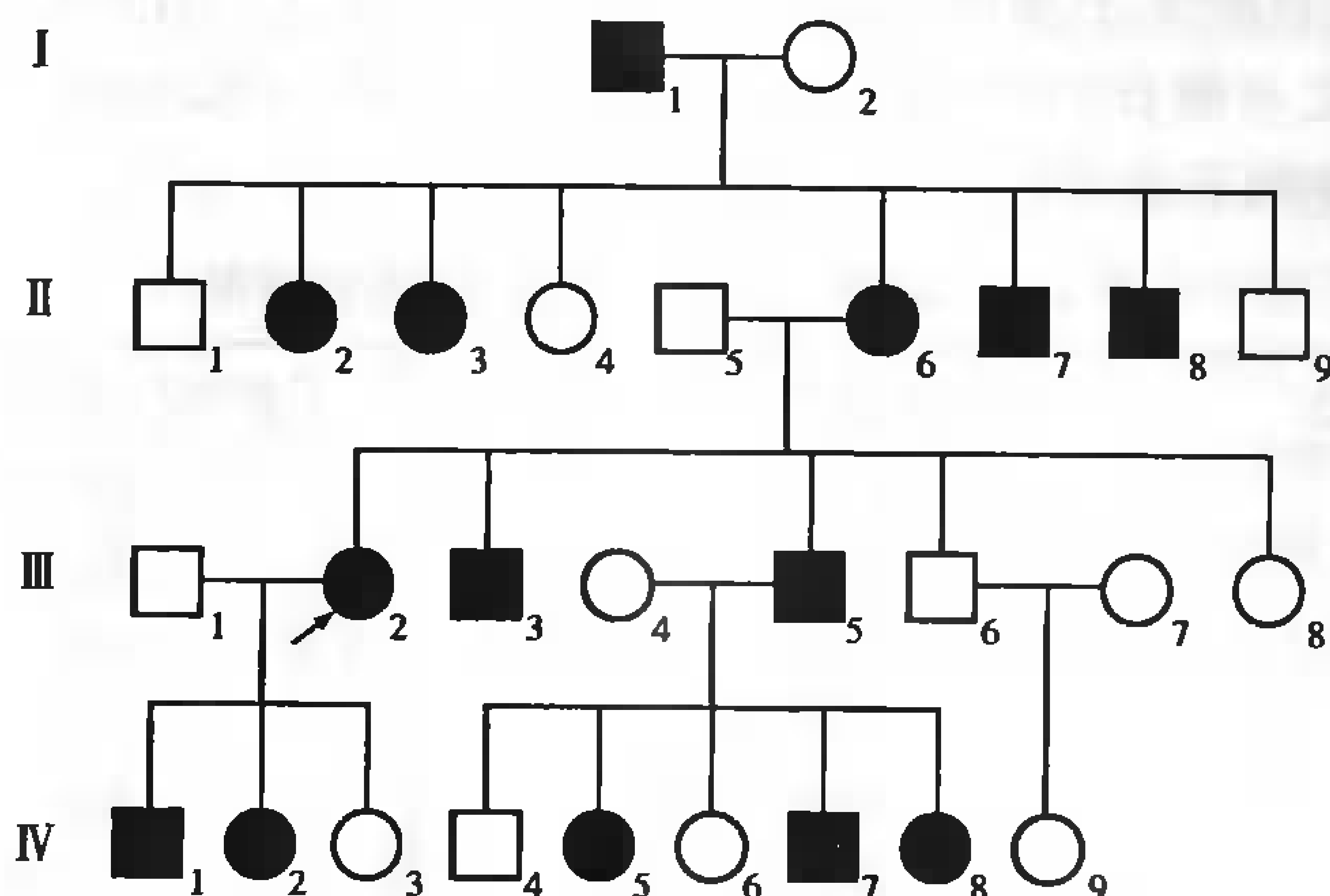


图 4-3 一例牙本质发育不全系谱

(二)不完全显性

不完全显性(incomplete dominance)是指杂合子(Aa)的表型介于纯合显性和纯合隐性之间,又称半显性(semidominance)。在不完全显性情况下,2个杂合子(Aa)婚配,子代的表型比例不是3:1,而是1:2:1。

人类对苯硫脲(PTC)的尝味能力就表现为不完全显性。PTC是一种白色结晶物,因含有 N—C═S 基团而具有苦涩味。PTC尝味能力受显性基因T的控制,显性纯合子(TT)能尝出浓度在1/750 000~1/3 000 000的PTC溶液的苦涩味,称为PTC尝味者;杂合子(Tt)能尝出浓度在1/50 000左右的PTC溶液的苦涩味,为杂合尝味者;而隐性纯合子(tt)只能尝出浓度大于1/24 000的PTC溶液的苦涩味,称为PTC味盲,有的味盲个体甚至连PTC结晶物的苦涩味也不能尝出。

当纯合显性尝味者(TT)与纯合隐性味盲(tt)个体婚配时,其子女都是杂合尝味者;当2个杂合尝味者个体婚配时,其子女中有1/4为纯合显性尝味者,1/2为杂合的尝味者,1/4为纯合隐性的味盲。已知味盲个体易患结节性甲状腺肿,因此,临床上可检测PTC尝味能力,作为结节性甲状腺肿的辅助诊断指标。

软骨发育不全、家族性高胆固醇血症及β地中海贫血等遗传病都表现为不完全显性。

(三)共显性

共显性(codominance)是指当等位基因处于杂合状态时,彼此间没有显性和隐性的区别,两者的作用都完全得到表现。

人类大多数血型系统的遗传、HLA系统的遗传均呈现共显性。人类的血型是由细胞膜表面的抗原物质决定的,而抗原物质又受基因的控制。ABO血型系统受控于9q34上的I^A、I^B和i一组复等位基因(multiple alleles)。复等位基因是指在群体中,一对特定的基因位点上有2个以上等位基因存在。I^A编码A抗原,I^B编码B抗原,i为无效基因,无基因产物。I^A、I^B对i是显性,I^A和I^B之间是共显性关系。因此,在人群中,这一组等位基因共有6种基因型,4种血型:A型($I^A I^A$、$I^A i$)、B型($I^B I^B$、$I^B i$)、O型(i i)及AB型($I^A I^B$)。其中,AB型者红细胞膜表面既有A抗原也有B抗原存在,表现为共显性。

根据显性遗传的规律，只要已知双亲的血型就可以推测出子女可能有和不可能有的血型；或已知双亲之一和子女的血型，也可以推测另一双亲可能有和不可能有的血型（表4－1）。这在法医学的亲子鉴定上有一定作用。

表4－1　双亲与子女ABO血型遗传的规律

双亲的血型	子女中可能有的血型	子女中不可能有的血型
A×A	A、O	AB、B
A×O	A、O	AB、B
A×B	A、B、AB、O	—
A×AB	A、B、AB	O
B×B	B、O	A、AB
B×O	B、O	A、AB
B×AB	A、B、AB	O
AB×O	A、B	AB、O
AB×AB	A、B、AB	O
O×O	O	AB、A、B

（四）不规则显性

由于某些因素的影响，使杂合子显性致病基因的作用没能表达，或者表达的程度有差异，这种现象称为不规则显性（irregular dominance）。影响显性基因表达的因素可以是遗传因素（主要是修饰基因对主基因的作用），也可以是环境因素。

显性基因的表达与否用外显率（penetrance）来衡量，外显率是指在带有显性致病基因的全部个体中，表现出相应症状的个体所占的百分率。完全外显时，外显率为100%，这种情况很少见；外显率低于100%时称为不完全外显。通常70%～80%的外显率就算比较高了，低的只有20%～30%。

例如，轴后型多指症就表现为不规则显性，多指症分轴前型和轴后型。轴前型多指患者的额外指位于大拇指的外侧，而轴后型多指患者的额外指位于小指的外侧。图4－4是一例轴后型多指症系谱，先证者$Ⅲ_2$的子女中$Ⅳ_1$有多指，$Ⅲ_2$的双亲$Ⅱ_3$和$Ⅱ_4$表型正常，但$Ⅲ_2$的伯父$Ⅱ_2$是多指患者，$Ⅱ_2$的母亲$Ⅰ_2$也是多指患者。根据系谱可知，$Ⅲ_2$的多指基因来源于他的父亲$Ⅱ_3$，$Ⅱ_3$带有多指基因，但可能由于修饰基因或（和）环境因素的影响而未表现出多指，称为未外显，并非是新发生的基因突变。$Ⅱ_3$虽未发病，但仍能将多指基因传递给他的子女，其子女患多指的可能性仍为1/2。

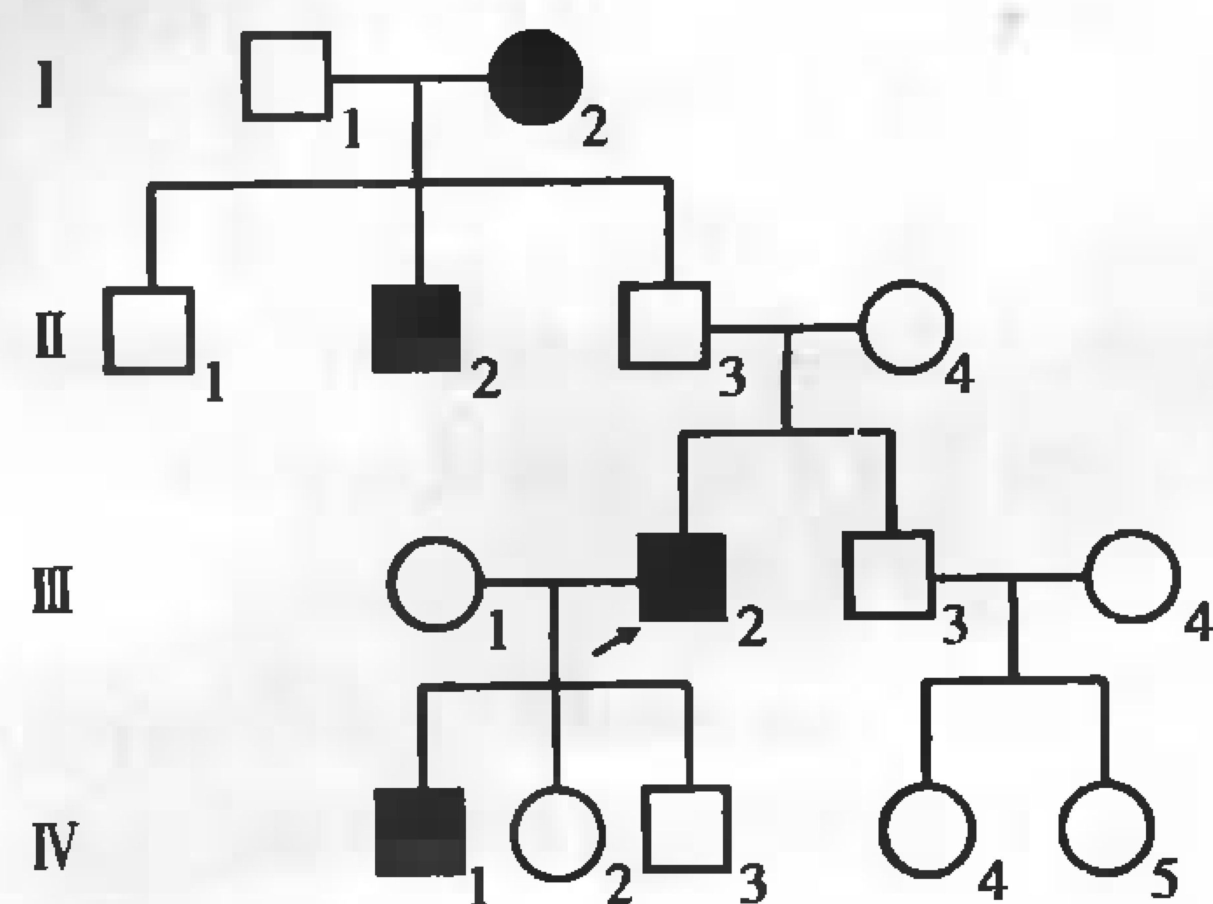


图4－4　一例轴后型多指症系谱

在不规则显性情况下，不同的杂合子患者显性基因的表达程度有差异，即具有相同基因型（Aa）的个体可能会表现出轻重不同的表型，这种个体间显性基因表达程度的差异称为表现度（expressivity）。

例如，Marfan 综合征(Marfan syndrome)是一种少见的结缔组织遗传病，患者的主要受累器官为骨骼、心血管系统和眼。临床表现为：身体瘦高、四肢细长，手指如蜘蛛样，颅骨细长，硬腭高拱，常见鸡胸或漏斗胸，常伴有韧带松弛及脊柱侧凸。眼部典型损害为晶状体脱位等。该病患者 60% ~ 80% 有心血管疾病，最常见是二尖瓣的功能障碍，心血管畸形常引起患者过早死亡。

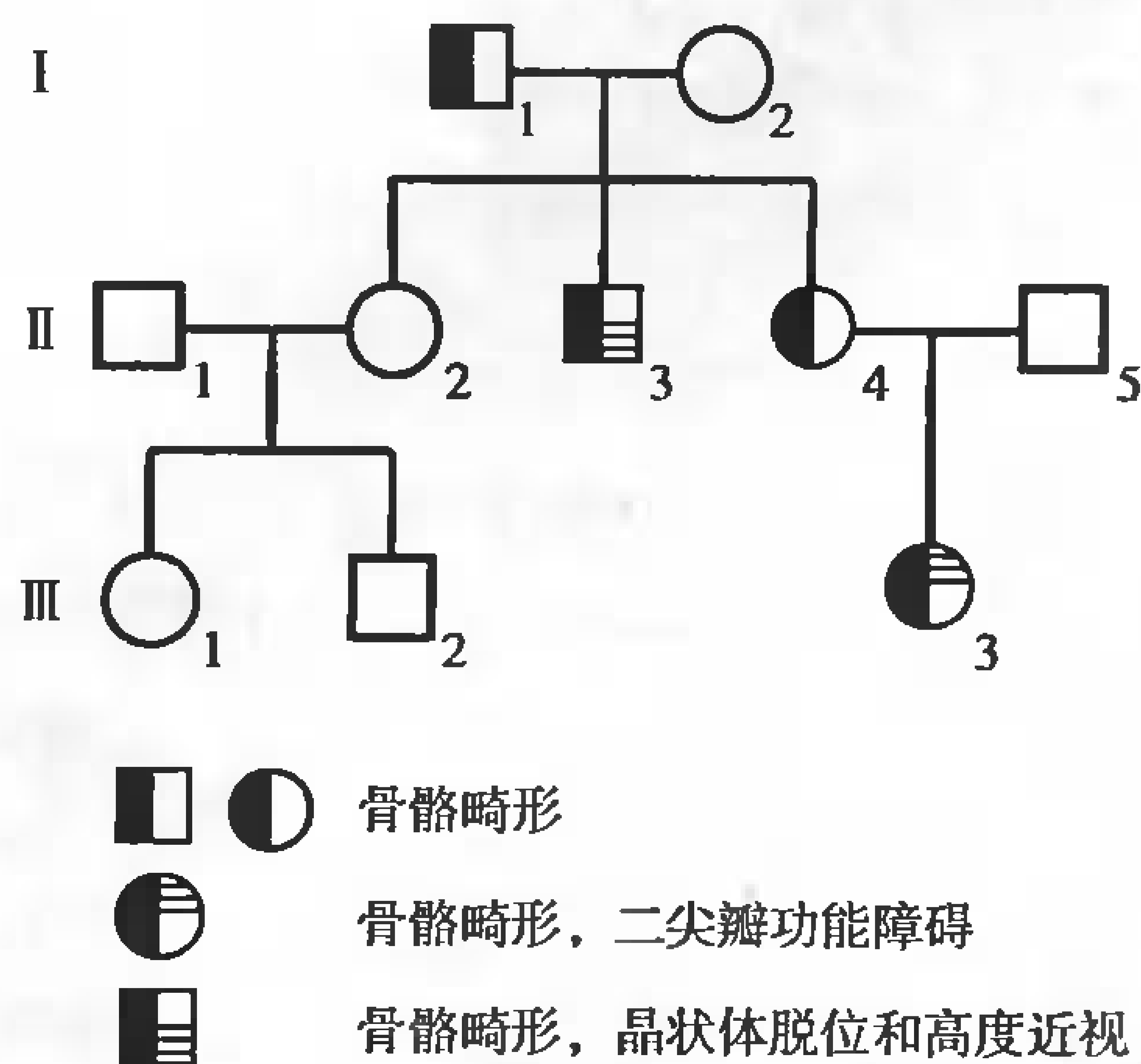


图 4－5　一例 Marfan 综合征系谱

本综合征的重型患者可有骨骼、眼、心血管系统的严重损害；轻型患者则只有少数器官轻度损伤。同一家系(图 4－5)中的不同患者可呈现出不同器官、不同程度损害，即表现度差异。

外显率与表现度是两个不同的概念，切不可混淆。其根本的区别在于前者阐明了基因表达与否，是"质"的概念；而后者要说明的是在表达前提下的表现程度差异，是"量"的概念。

(五)延迟显性

延迟显性(delayed dominance)是指有些 AD 病，在生命的早期，杂合子携带的致病基因并不表达或虽表达但尚不足以引起明显的临床表现，只有在达到一定年龄，致病基因才表达或才充分表达并表现出疾病的现象。例如，在亨廷顿舞蹈病，多数患者(杂合子)在 35 ~ 40 岁时发病，少数患者在十几岁或 60 多岁发病。

二、常染色体隐性遗传

常染色体上的隐性基因控制的疾病或性状的遗传称为常染色体隐性遗传(autosomal recessive inheritance，AR)。在 AR 病中，只有隐性纯合子(bb)才发病；杂合子(Bb)虽然带有致病基因，但隐性致病基因的作用被显性正常基因掩盖，表型与正常人相同；杂合子却可将隐性致病基因遗传给后代。这种带有致病基因但表型正常的个体称为携带者(carrier)。常见的 AR 病包括先天性聋哑、苯丙酮尿症、白化病等。

群体中 AR 病致病基因的频率很低，一般为 0.01 ~ 0.001。患者的双亲往往不发病，但一定是致病基因携带者。2 个携带者(Bb × Bb)婚配，子女的基因型有 3 种：BB、Bb 和 bb，其比例为 1:2:1。其中，纯合隐性患者占 1/4，表型正常个体占 3/4，但表型正常个体有 2/3 的可能性为携带者。

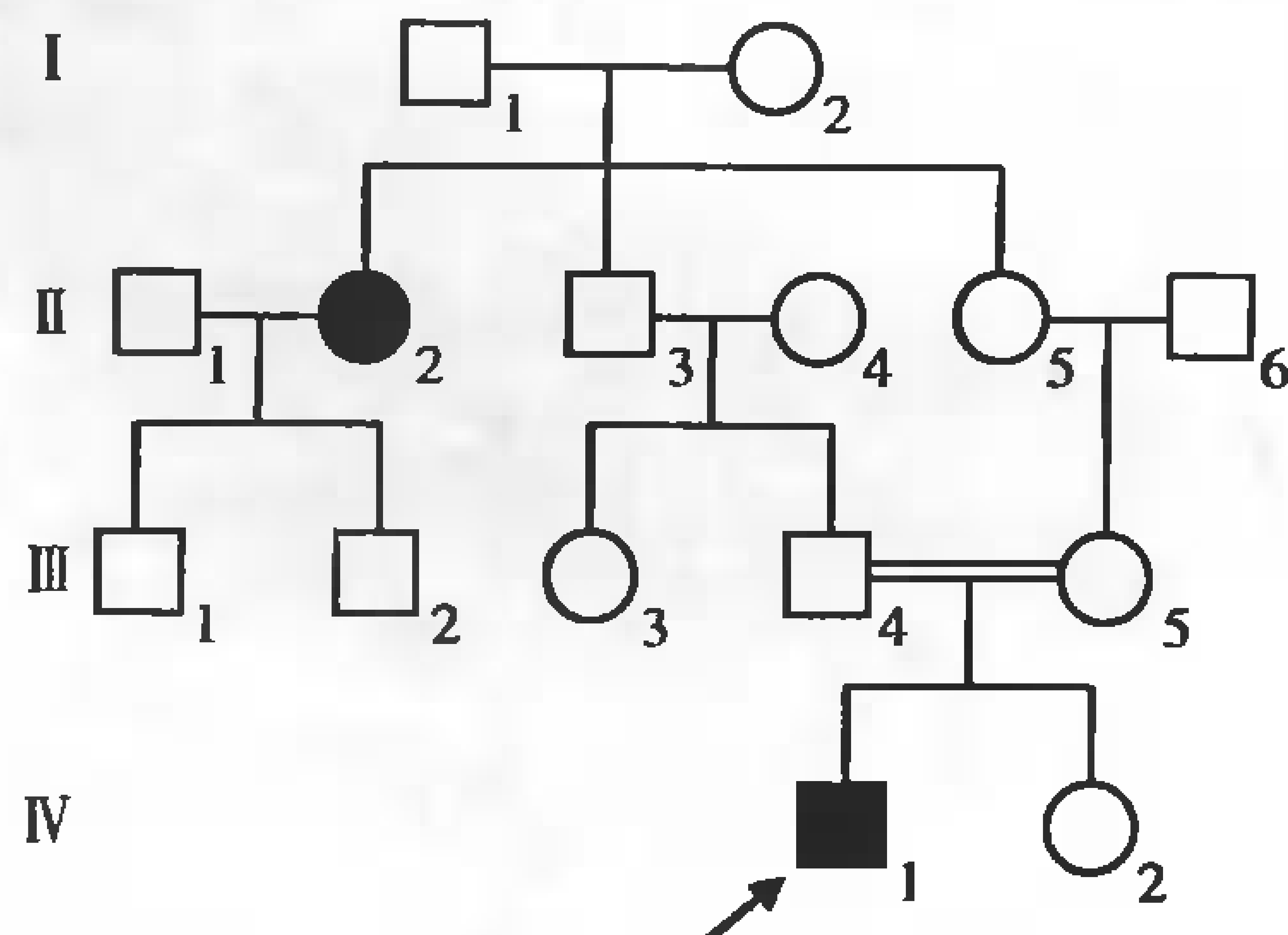


图 4－6　一例白化病系谱

图 4－6 是一白化病系谱，该病患者皮肤

毛发呈白色,虹膜淡灰色,畏光,眼球震颤。该病是由于编码酪氨酸酶的基因(15q11～q12)突变,导致酪氨酸酶缺陷,不能产生黑色素所致。系谱中先证者$Ⅳ_1$的双亲$Ⅲ_4$和$Ⅲ_5$表型正常,但他们却生出了白化病患儿,这说明他们都是肯定携带者(obligatory carrier)。根据孟德尔定律,他们所生的每个孩子都有1/4的可能性是白化病患儿。

根据以上白化病家系,可见AR病系谱具有以下主要特点:①患者双亲往往表型正常,但都是致病基因携带者;②患者同胞患病可能性为1/4,男女发病机会相等;③患者子女一般不发病,系谱中看不到连代传递现象,病例往往是散发的;④近亲婚配可使子女发病风险明显增高。

(一)患者同胞的发病风险

理论上,AR病患者同胞患病风险为1/4,但实际调查的结果往往大于1/4,这是由于选择偏倚造成的。在父母均为同一致病基因携带者的家庭中,子女中有患者的家庭被统计,无患病子女的家庭将被漏计。①在只生1个孩子的家庭中,不发病(3/4)的孩子被漏计,发病(1/4)的孩子被统计。因此,将1/4的发病率误计为100%。②在生2个孩子的家庭中,2个孩子都患病的可能性为$(1/4)^2=1/16$,被统计;2个孩子1个患病1个正常的可能性为$2\times(1/4\times3/4)=6/16$,被统计;2个孩子都正常的可能性为$(3/4)^2=9/16$,被漏计。同理,生$n$个孩子的家庭中将有$(3/4)^n$个家庭被漏计。所以,在实际调查中,患者比例往往超过1/4。因此,在计算AR病家系中患者同胞的发病风险时,需要校正统计结果。常用的校正方法是Weinberg先证者法。校正公式为:

$$C=\frac{\sum a(r-1)}{\sum a(s-1)}$$

公式中,C为校正值,即患者同胞的实际发病风险;a为先证者人数,恒等于1;r为包括先证者在内的同胞中受累人数;s为同胞人数。此公式校正偏倚的基本原理是,将先证者从统计中除去,只算先证者同胞的发病比例,先证者只起指认2个携带者婚姻的作用。例如,表4-2是一项对苯丙酮尿症家庭的调查结果,在11个家庭的23名同胞中,有患者14人,患者同胞发病风险为14/23≈0.609。应用校正公式计算,$C=3/12=1/4$,符合AR病同胞发病风险的理论比例。

表4-2　苯丙酮尿症患者同胞的发病比例校正值

s	r	a	$a(r-1)$	$a(s-1)$
1	1	1	0	0
1	1	1	0	0
1	1	1	0	0
1	1	1	0	0
2	1	1	0	1(2-1)
2	1	1	0	1(2-1)
2	2	1	1(2-1)	1(2-1)
3	1	1	0	1(3-1)
3	1	1	0	1(3-1)
3	2	1	1(2-1)	1(3-1)
4	2	1	1(2-1)	1(4-1)
23	14	11	3	12

(二)近亲婚配增高子代的发病风险

1.近亲与近亲婚配　近亲(consanguinity)是指3~4代以内有共同祖先的个体,近亲个体之间的婚配称近亲婚配(consanguineous marriage)。由于近亲之间具有共同的祖先,故带有同一基因的可能性相对较大。近亲婚配中的一个是某种致病基因的携带者时,另一个也是携带者的可能性远高于群体中的携带者频率。因此,他(她)们所生子女成为隐性纯合子的机会比随机婚配时要高得多。即近亲婚配所生子女AR病发病风险增高。

2.亲缘系数　亲缘系数(coefficient of relationship)是指2个有共同祖先的个体在某一基因位点具有相同等位基因的概率。

设一对双亲P_1和P_2某一基因位点的等位基因分别为A_1A_2、A_3A_4,他(她)们生有2个子女:B_1和B_2。亲子之间基因相同的概率:父亲或母亲都将自己一半的基因传给子女,即B_1的基因型有4种可能性,即A_1A_3、A_1A_4、A_2A_3和A_2A_4。不管是哪一种,B_1与亲代都有1/2的基因相同,即亲子之间基因相同的概率为1/2。同胞之间基因相同的概率:在4种基因型中,有A_1基因(A_1A_3和A_1A_4)的占1/2,有A_2、A_3和A_4基因的也各占1/2。因此,B_1与B_2(同胞)之间基因相同的概率为1/2。即亲子之间、同胞兄弟姐妹之间的亲缘系数为1/2。

值得注意的是,亲子之间亲缘系数1/2是绝对的,即亲子之间必有1/2基因相同。同胞之间亲缘系数1/2仅仅是概率估计,即某一基因位点上有1/2的可能性基因相同,而实际上,也存在同胞之间某一基因位点上基因完全相同,或完全不同(A_1A_3与A_2A_4)的情况。

亲缘系数为1/2的亲属称为一级亲属(first degree relative)。同理,一个人与其祖辈、孙辈之间,与叔、伯、姑、舅、姨、之间的亲缘系数为1/4,称为二级亲属(second degree relative);与曾祖辈、曾孙辈、祖辈的同胞、同胞的孙辈、表兄妹及堂兄妹等之间的亲缘系数为1/8,称为三级亲属(third degree relative)。

3.近亲婚配与子女的AR病发病风险　图4-7是一AR病系谱,设群体中该病致病基因频率为0.01,根据遗传平衡定律(参见第九章　群体遗传学)可知,正常基因频率为$1-0.01=0.99$,携带者频率为$2\times0.01\times0.99=0.0198$,约1/50。系谱中Ⅱ$_1$是患者,说明Ⅰ$_1$和Ⅰ$_2$都是携带者,Ⅱ$_2$和Ⅱ$_4$是携带者的可能性均为2/3。所以,Ⅲ$_1$和Ⅲ$_2$是携带者的频率均为$2/3\times1/2=1/3$。在AR病中,两个携带者婚配,子代的发病比例为1/4,所以Ⅳ$_1$发病的可能性为$1/3\times1/3\times1/4=1/36$。如果Ⅲ$_1$不与Ⅲ$_2$婚配,而是在群体中随机婚配,所生子女的发病风险将会明显降低。因为群体中携带者频率为1/50,所以子女的发病风险为$1/50\times1/3\times1/4=1/600$。与表兄妹婚配相比,随机婚配子女的发病风险下降了近17倍。

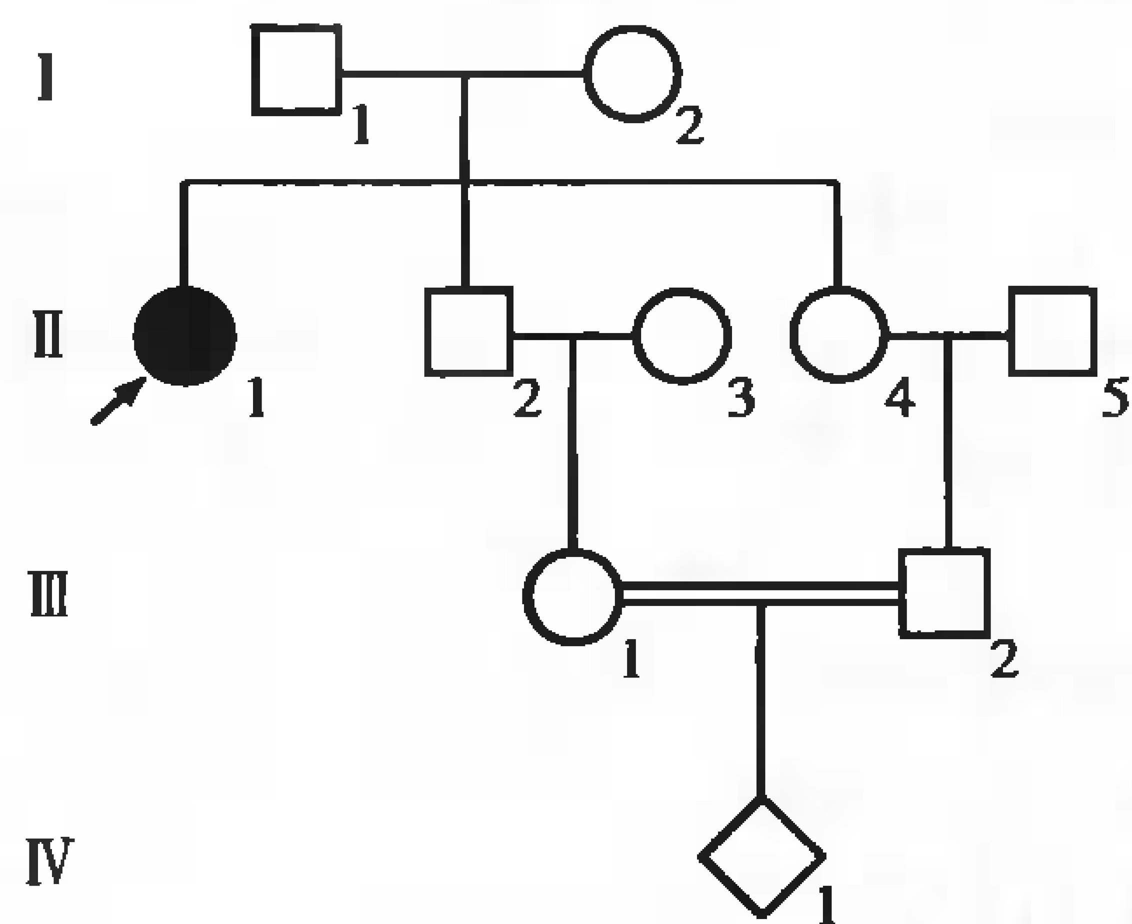


图4-7　一例AR病系谱

假如上述家系中Ⅱ$_1$是正常人,那么Ⅲ$_1$、Ⅲ$_2$是携带者的频率均为1/50,如果他们均随机婚配,所生子女发病风险为$1/50\times1/50\times1/4=1/10\,000$。如果Ⅲ$_1$和Ⅲ$_2$婚配,其中一人是携带者的可能性仍为1/50,但另一人是他的三级亲属,基因相同的可能性为1/8,

所以，所生子女发病的可能性为 1/50 × 1/8 × 1/4 = 1/1 600。与随机婚配相比，子女的发病风险增加了 6 倍多。如果群体致病基因频率为 0.001，群体中携带者频率则为 1/500，随机婚配所生子女的发病风险为 1/500 × 1/500 × 1/4 = 1/1 000 000；而 $Ⅲ_1$ 和 $Ⅲ_2$ 婚配所生子女的发病风险则为 1/500 × 1/8 × 1/4 = 1/16 000，两相比较，发病风险增加了 60 多倍。这说明，随着致病基因频率的降低，产生 AR 病的绝对风险率也随之降低；但与随机婚配相比，近亲婚配使子代罹患 AR 病的相对危险率更高。即越是罕见的 AR 病，近亲婚配的危害性越大。

三、X 连锁遗传

X 染色体上的基因所控制的性状或疾病的遗传，称为 X 连锁遗传（X – linked inheritance）。根据基因性质的不同，X 连锁遗传分为 X 连锁显性遗传和 X 连锁隐性遗传两类。X 连锁基因的遗传不同于常染色体上基因的遗传，男性的 X 连锁基因只能随 X 染色体从母亲传来，将来只能传给女儿，不可能传给儿子，这种遗传称为交叉遗传（criss – cross inheritance）。

（一）X 连锁显性遗传

X 染色体上的显性基因所控制的性状或疾病的遗传，称为 X 连锁显性遗传（X – linked dominant inheritance，XD）。女性有 2 条 X 染色体，显性纯合子（XAXA）和杂合子（XAXa）都会患病，但群体中致病基因的频率通常很低，女性为显性纯合患者的可能性极小。因此，人群中的女患者一般为杂合子。男性只有 1 条 X 染色体，其上带有致病基因（X^AY）即患病，不带致病基因（X^aY）不发病。所以在 X 连锁显性遗传病中，女性的发病率是男性的 2 倍，但女患者的病情往往较男患者轻。

男性患者（X^AY）与正常女性婚配（X^aX^a），女儿（X^AX^a）都患病，儿子（X^aY）都正常；女患者（X^AX^a）与正常男性（X^aY）婚配，女儿（X^AX^a 或 X^aX^a）和儿子（X^AY 或 X^aY）均有 1/2 的可能患病。

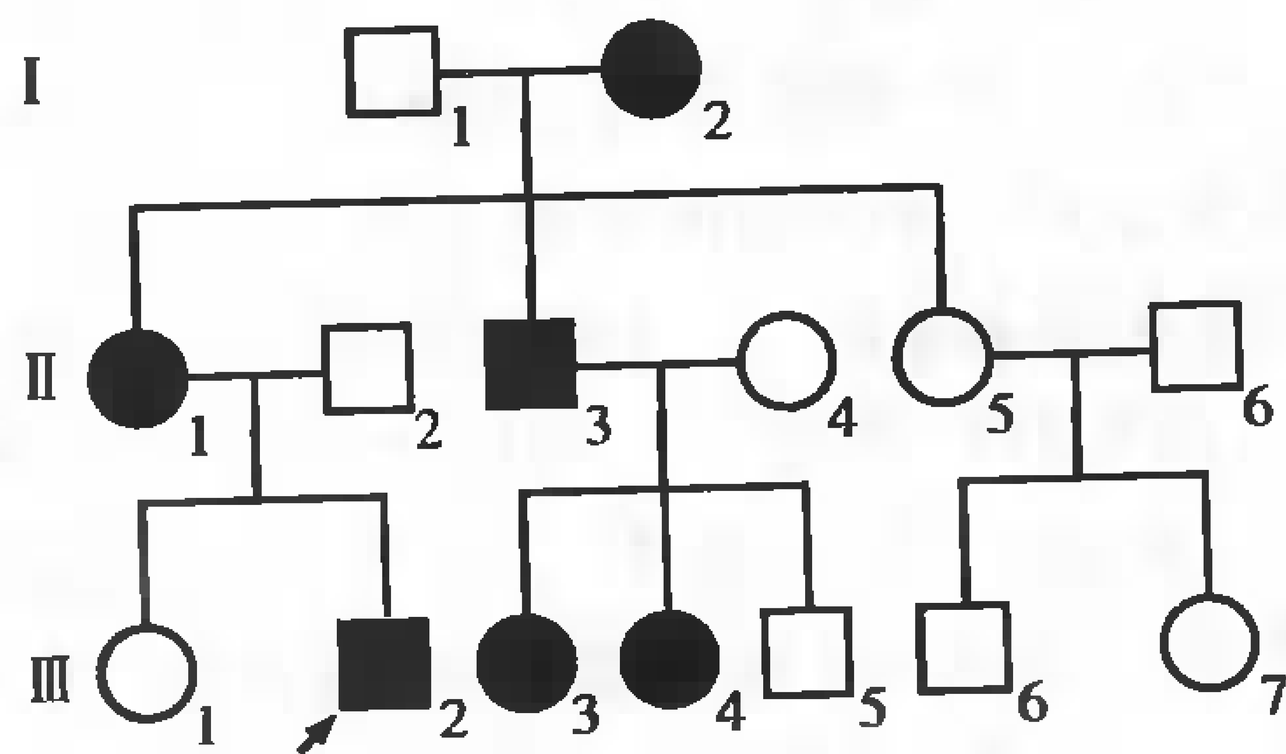


图 4 – 8　一例抗维生素 D 佝偻病系谱

常见的 XD 病有抗维生素 D 佝偻病（vitamin D resistant rickets）和遗传性肾炎等。图 4 – 8 是抗维生素 D 佝偻病的系谱，该病致病基因位于 Xp22。患者小肠对钙磷的吸收及肾小管对磷酸盐的重吸收障碍，导致血磷降低、尿磷增加，骨质钙化不全；临床上有 O 形腿、X 形腿、鸡胸等骨骼发育畸形和生长缓慢等佝偻病表现。患者服用常规剂量的维生素 D 无效，只有大剂量维生素 D 和磷的补充才能见效，故称为抗维生素 D 佝偻病。

根据抗维生素 D 佝偻病系谱，可见 X 连锁显性遗传病具有以下主要特点：①人群中女患者多于男患者，但女患者病情往往较男患者轻；②男患者的母亲是患者，女患者的双亲之一是患者；③男患者的女儿都发病，儿子都正常；女患者的子女均有 1/2 发病；④系谱中可见连续遗传。

(二)X连锁隐性遗传

X染色体上的隐性基因所控制的性状或疾病的遗传,称为X连锁隐性遗传(X - linked recessive inheritance, XR)。在XR中,由于男性只有1条X染色体,故只要在X染色体上带有致病基因(Xh)就将患病(XhY);女性有2条X染色体,1条X染色体上有致病基因不患病,但为X连锁隐性致病基因携带者(XHXh),只有2条X染色体上都有致病基因才患病(XhXh)。因此,人群中男患者多于女患者。

男患者(X^hY)与正常女性(X^HX^H)婚配,儿子(X^HY)都正常,女儿(X^HX^h)都是携带者。

男患者(X^hY)与女性携带者(X^HX^h)婚配,儿子(X^HY或X^hY)1/2正常,1/2患病;女儿(X^hX^h或X^HX^h)1/2患病,1/2是携带者。

正常男性(X^HY)与女性携带者(X^HX^h)婚配,儿子(X^HY或X^hY)1/2正常,1/2患病;女儿(X^HX^H或X^HX^h)1/2正常,1/2是携带者。

常见XR病有:假性肥大性肌营养不良症(pseudohypertrophic muscular dystrophy)、红绿色盲(protanopia and deuteranopia)及甲型血友病(hemophilia A)等。

甲型血友病是一种典型的X连锁隐性遗传病(图4-9),致病基因定位于Xq28。由于缺乏凝血因子Ⅷ,患者出现凝血障碍,在皮下、肌肉内反复出血,可形成瘀斑或血肿;在关节腔出血可致关节畸形;严重者可因颅内出血而死亡。

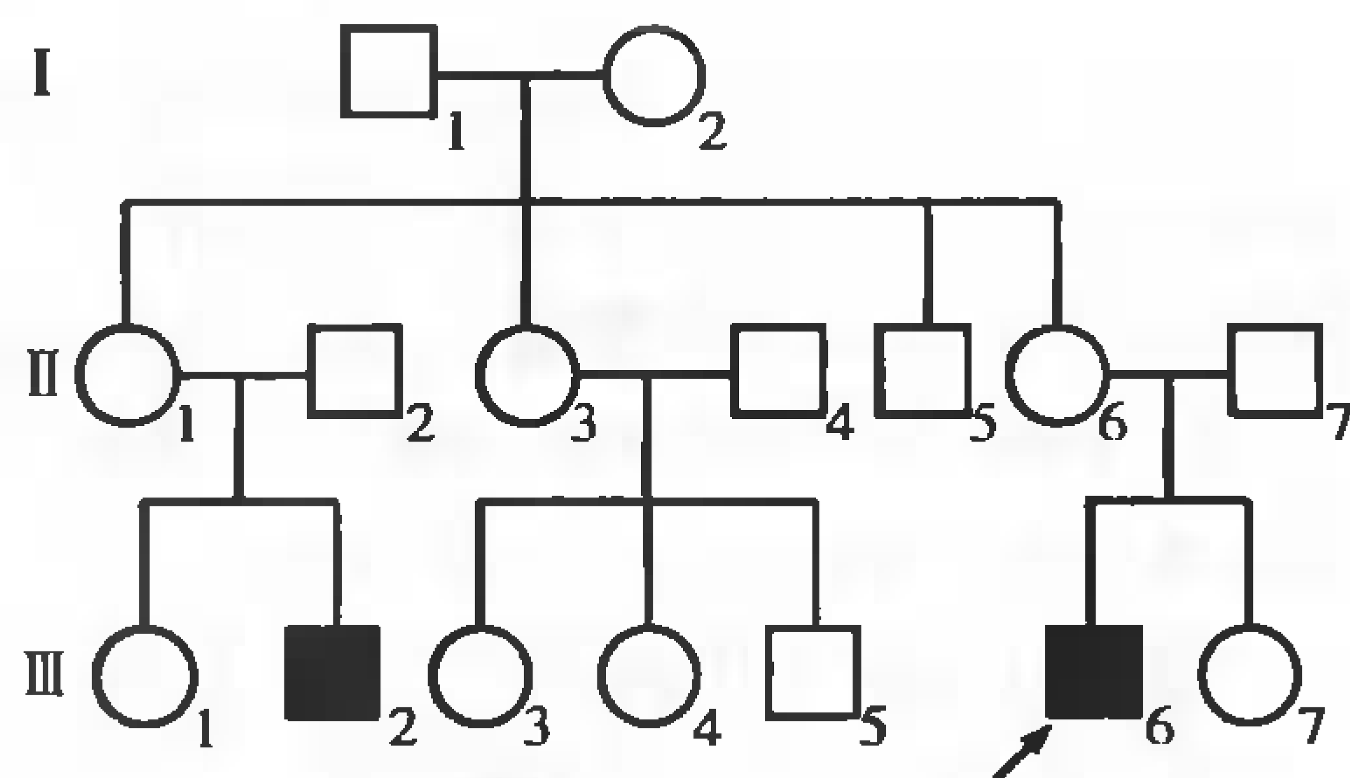


图4-9　一例甲型血友病系谱

系谱中,患者$Ⅲ_2$和$Ⅲ_6$的致病基因分别来自她们的母亲$Ⅱ_1$和$Ⅱ_6$,故$Ⅱ_1$和$Ⅱ_6$都是血友病基因携带者,而她们的致病基因均来自$Ⅰ_2$,因为$Ⅰ_1$如有致病基因必然是患者,$Ⅱ_5$未发病,说明遗传了母亲的正常基因;$Ⅱ_3$有1/2的可能性是携带者,$Ⅲ_1$和$Ⅲ_7$也均有1/2的可能性是携带者。

X连锁隐性遗传病具有以下主要特点:①人群中男患者远多于女患者,系谱中往往只见到男患者;②双亲无病,儿子可能发病,女儿不会发病;儿子的致病基因来自携带者母亲;③男患者的兄弟、舅父、姨表兄弟、外甥可能是患者;④由于大多数XR病有致死效应,故外祖父、外孙发病少见;外祖父发病时,舅父不会发病。

四、Y连锁遗传

Y染色体上的基因控制的性状或疾病的遗传称为Y连锁遗传(Y - linked inheritance)。由于Y染色体只能从父亲传递给儿子,再由儿子传递给孙子,所以Y连锁遗传又叫全男性遗传或限雄性遗传(holandric inheritance)。

性别决定基因(SRY)位于Yp11.2,长约35kb,只有1个外显子,该基因突变、缺失、易位可导致两性畸形或性逆转综合征。由于带有SRY突变基因个体不能生育,故该病不会传给后代。无精子症因子(azoospermia factor, AZF)基因位于Yq11.23,该基因异常时导致无精症或少精症。

外耳道多毛受 Y 染色体上的外耳道多毛基因控制，受累男性青春期后在外耳道长出 2～3 cm 长、成丛的黑色硬毛，图 4－10 是一外耳道多毛家系，可见该家系中男性均有此性状，女性均无此性状。

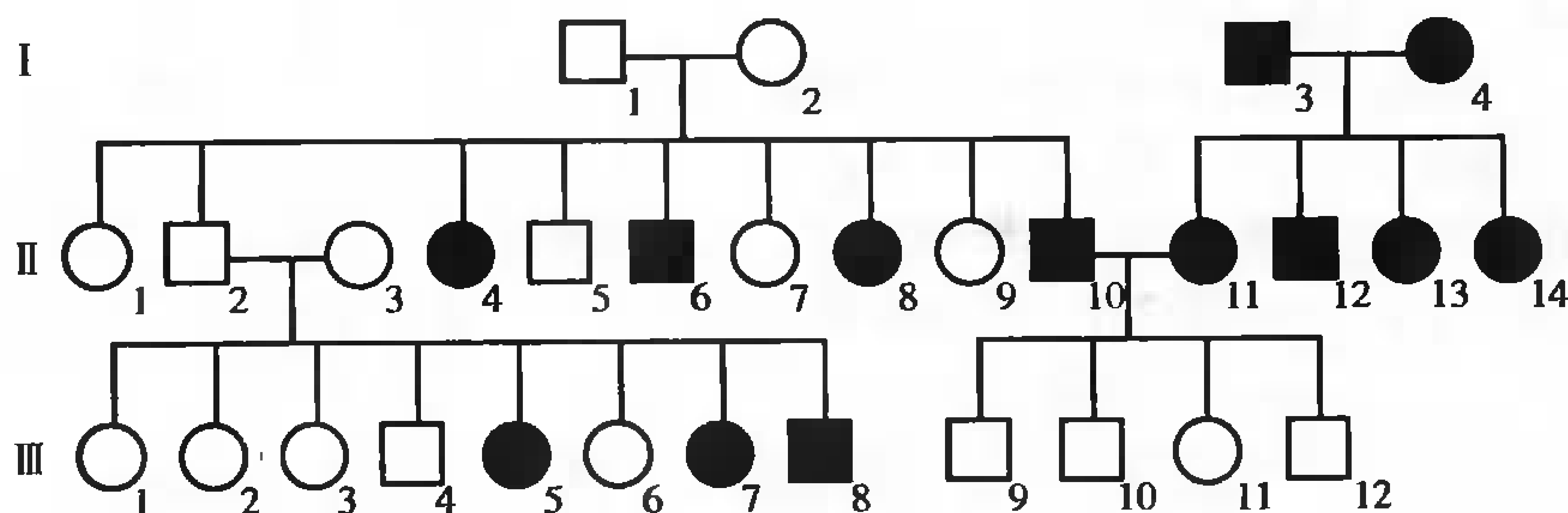


图 4－10　一例外耳道多毛系谱

五、2 种单基因性状的独立传递

如果一个家系中同时出现 2 种性状或疾病，并且决定这 2 种性状的基因位于非同源染色体上，那么这 2 种性状的传递符合自由组合律。例如，父亲是并指患者，母亲表型正常，生育了 1 个先天性聋哑患儿。这对夫妇再生孩子的表型如何?

设并指(AD)和先天性聋哑(AR)的致病基因分别为 D 和 S，则患儿及患儿的母亲和父亲的基因型分别为 ddss、ddSs、DdSs。

根据自由组合律，母亲可以产生 dS 和 ds 2 种类型的卵子，父亲可以产生 DS、Ds、dS 及 ds 4 种类型的精子。精卵随机组合，子代可能的表型：并指伴先天性聋哑的概率为 $1/2 \times 1/4 = 1/8$；只有并指而不患先天性聋哑的概率为 $1/2 \times 3/4 = 3/8$；无并指只有先天性聋哑的概率为 $1/2 \times 1/ = 1/8$；表型正常的概率为 $1/2 \times 3/4 = 3/8$。

六、2 种单基因性状的联合传递

当决定 2 种性状或疾病的基因位于同一条染色体上时，这 2 种性状的传递符合连锁交换规律。例如，红绿色盲与甲型血友病的致病基因都位于 Xq28，在传递时彼此连锁。假如这两个基因之间的交换率为 10%。某一家庭父亲患红绿色盲，母亲表型正常，已生了 1 个患红绿色盲的女儿和 1 个患甲型血友病的儿子。这对夫妻如再次生育，子代发病风险如何?

设红绿色盲基因为 X^b，甲型血友病基因为 X^h。家庭中儿子的基因型为 $X^{Bh}Y$，他的 XBh 染色体来自母亲；色盲女儿的基因型为 $X^{bH}X^{bH}$，其中一条 X^{bH} 染色体来自父亲($X^{bH}Y$)，另一条 X^{bH} 染色体来自于母亲($X^{Bh}X^{bH}$)。

由于存在 10% 的交换母亲可以产生 4 种类型的卵子：非交换型 2 种，X^{Bh} 和 X^{bH}，各占 45%；交换型 2 种，X^{BH} 和 X^{bh}，各占 5%。而父亲只产生 XbH 和 Y 2 种类型的精子。因此，这对夫妻所生的子女中：女儿表型正常或患红绿色盲的可能性均为 50%；儿子患血友病或红绿色盲的可能性均为 45%，既患血友病又患色盲的可能性为 5%，两种病都不患的可能性也为 5%。

第二节　影响单基因遗传效应的因素

一、遗传背景

单基因遗传病的发病主要受控于一对主基因，但基因组中除决定某一性状的主基因外，还有许多其他影响疾病表现的基因，包括互补基因、上位基因、修饰基因及抑制基因等。互补基因是指非等位显性基因之间的作用对某一性状的表现是同等重要的，它们之间必须互补，才能显示其相应的效应。上位基因是指某对等位基因不仅有自己独立的效应，而且可影响或掩盖另一非等位基因的作用；若掩盖者为显性基因，称为显性上位基因；若掩盖者为隐性基因，称为隐性上位基因。修饰基因本身不产生任何表型效应，但它可影响其他等位基因的作用。抑制基因本身也没有独立的表型效应，但可抑制另一等位基因的作用。不同基因之间的相互作用共同决定了疾病的表现。可以认为，任何一个性状的形成，包括疾病的产生，都受许多基因的影响。一般将基因组中决定某一性状或疾病表现的主基因以外的其他基因称为这对基因的遗传背景（genetic background）或基因环境。具有同一种致病基因的不同个体，由于遗传背景不同，致病基因在不同个体的表型效应就可能不同。例如，在轴后型多指症和 Marfan 综合征家系中不同个体之间外显率和表现度的差异形成机制就在于此。

二、基因多效性

基因多效性（gene pleiotropy）是指一个或一对基因可以产生多种表型效应。个体发育过程中，很多生理生化过程都是相互联系、相互依赖的。一个或一对基因异常常常会在不同组织或个体发育的不同阶段引起一系列的生化代谢或组织结构的异常，从而呈现出疾病的多种表现。例如，在苯丙酮尿症（AR），由于编码苯丙氨酸羟化酶的基因突变，苯丙氨酸羟化酶缺陷，导致苯丙氨酸主要代谢途径受阻，苯丙氨酸代谢旁路增强，产生过多的苯丙酮酸及其衍生物从尿中排出，称为苯丙酮尿；而苯丙酮酸及其衍生物又可影响脑的发育，造成智力障碍；酪氨酸生成障碍，以致黑色素不能形成，出现“白化”表现。因此，虽然只是一对基因的突变，但却可出现多种异常表现。

三、遗传异质性

遗传异质性（genetic heterogeneity）是指表型相似而基因型不同的现象，又称为多因一效。遗传异质性分为等位基因异质性与位点异质性 2 种类型。

（一）等位基因异质性

等位基因异质性（allelic heterogeneity）是指同一基因位点发生不同的突变，导致同一疾病的不同患者具有不同的基因型，表型可能相似，也可能差异较大。例如，正常血红蛋白 HbA 的 β 珠蛋白基因第 6 位密码子为 GAG，如果 GAG 突变为 GUG，相应的珠蛋白多肽链第 6 位谷氨酸被缬氨酸取代，形成异常血红蛋白 HbS，临床上表现为镰形红细胞贫血症；如果 GAG 突变为 AAG，相应的珠蛋白多肽链第 6 位谷氨酸被赖氨酸取代，形成异常血红

蛋白 HbC，临床上表现为轻度溶血性贫血，可伴有肝脾肿大。

(二)位点异质性

位点异质性(locus heterogeneity)是指多个不同位点的基因作用于同一器官的发育，产生相同或相似的表型效应。而这些表型相似的遗传病可表现出相同或不同的遗传方式。例如，先天性聋哑的遗传方式有 AR、AD 和 XR 等，其中大多数是 AR。目前已经发现，常染色体隐性遗传的先天性聋哑，共有 35 个基因位点，这 35 个位点上任一基因处于纯合状态，均可导致先天性聋哑。如果一对聋哑夫妇的聋哑基因不在相同的基因位点上，所生子女均不聋哑(图 4－11)。

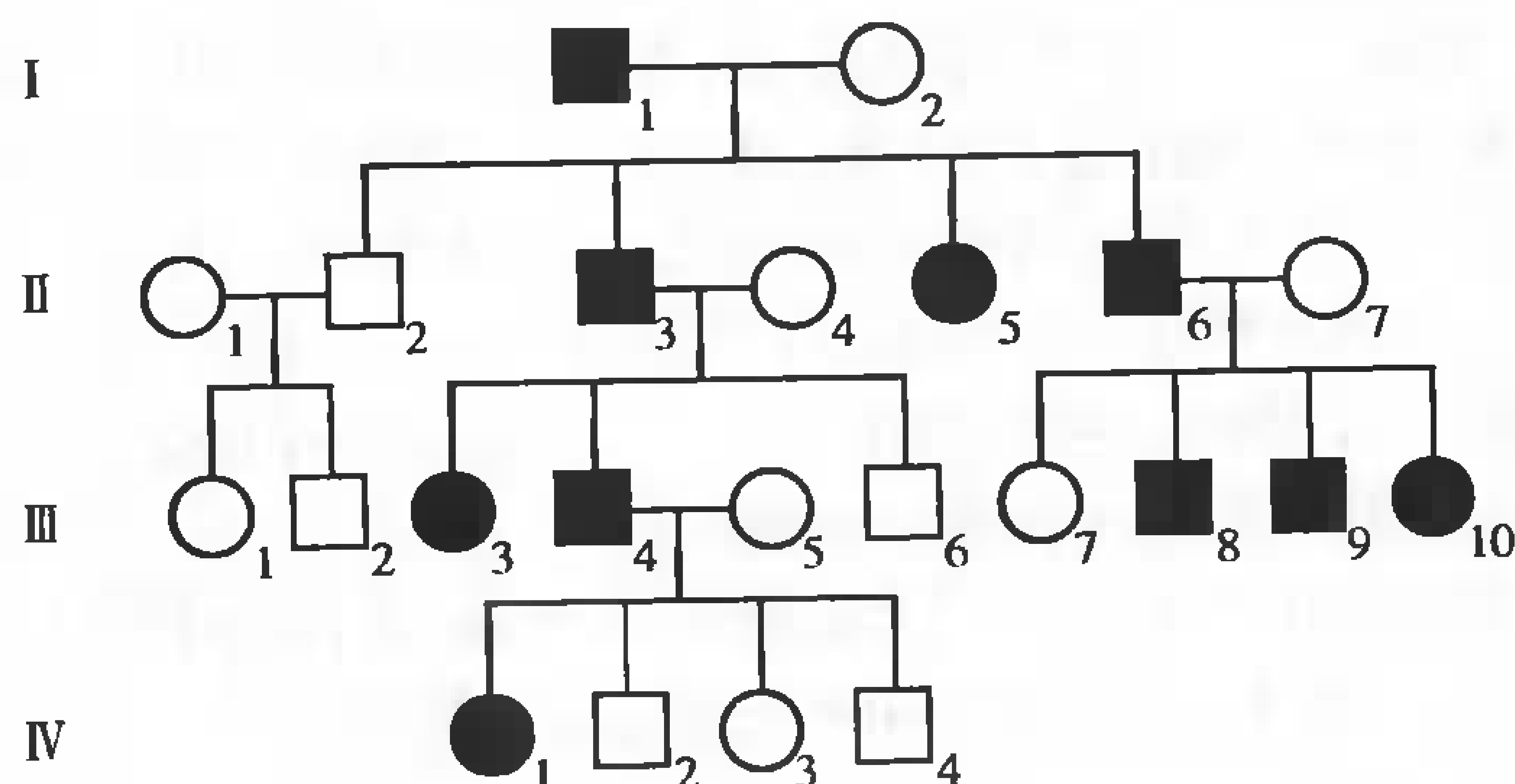


图 4－11 一例先天性聋哑系谱(示异质性)

环境因素所致的先天性聋哑约占 20%。风疹病毒可严重影响胎儿内耳的发育而致先天性聋哑。母亲在妊娠早期感染风疹病毒所致的先天性聋哑与常染色体隐性遗传的先天性聋哑具有相同的表型。这种环境因素作用使个体表现出与某一特定基因作用产生的表型相同或相似的现象称为拟表型(phenocopy)或表型模拟。拟表型是环境因素影响所致，并非生殖细胞中基因突变。因此，这种聋哑不遗传给后代。

由于遗传基础不同，具有遗传异质性的遗传病的遗传方式、发病年龄、病程进展、病情严重程度、预后以及复发风险等都可能不同。人群中遗传病病种的增多，除新发生的突变所致外，从已知具有遗传异质性的遗传病中分出了亚型也是重要原因。例如，血友病分为甲型、乙型及丙型，其中甲型和乙型为 X 连锁隐性遗传，丙型则为常染色体隐性遗传。

四、基因组印记

来自父方和母方的同源染色体或等位基因存在功能上的差异，传给子女时可形成不同的表型，这种现象称为基因组印记(genomic imprinting)或遗传印记(genetic imprinting)或亲代印记(parental imprinting)。

例如，亨廷顿舞蹈病的致病基因如果经母亲传递，子女的发病年龄与母亲的发病年龄相似；如果经父亲传递，则子女的发病年龄比父亲的发病年龄略有提前；在有的家系中，子女的发病年龄可以提前到 20 岁左右。基因组印记的机制比较复杂，初步研究提示，DNA 的甲基化是其重要的分子机制。精子和卵子中一些基因的甲基化程度不同，高度甲基化(被打上印记)的基因不表达或低表达；在胚胎发育过程中发生去甲基化时，这些基因开始

表达。基因组印记通常发生在配子形成期，印记持续存在个体的一生中。但它不是一种突变，在下一代配子形成时，旧的印记可以消除，并且可产生新的印记。因此，亨廷顿舞蹈病发病年龄提前的父源效应，经过一代传递即消失，早发型男性的后代仍为早发型，早发型女性的后代的发病年龄并不提前。

五、遗传早现

遗传早现(genetic anticipation)是指某些遗传病(主要是AD病)在世代传递过程中，发病年龄逐代提前、病情逐代加重的现象。例如，脊髓小脑性共济失调是一种AD病，发病年龄一般为35~40岁，早期表现为行走困难，站立摇摆不定，语言不清；晚期则有下肢瘫痪。在图4-12家系中，$Ⅰ_1$ 39岁开始发病，$Ⅱ_3$ 38岁开始发病，$Ⅲ_1$ 23岁就已瘫痪。在该病的许多家系中，都可看到这种遗传早现现象。

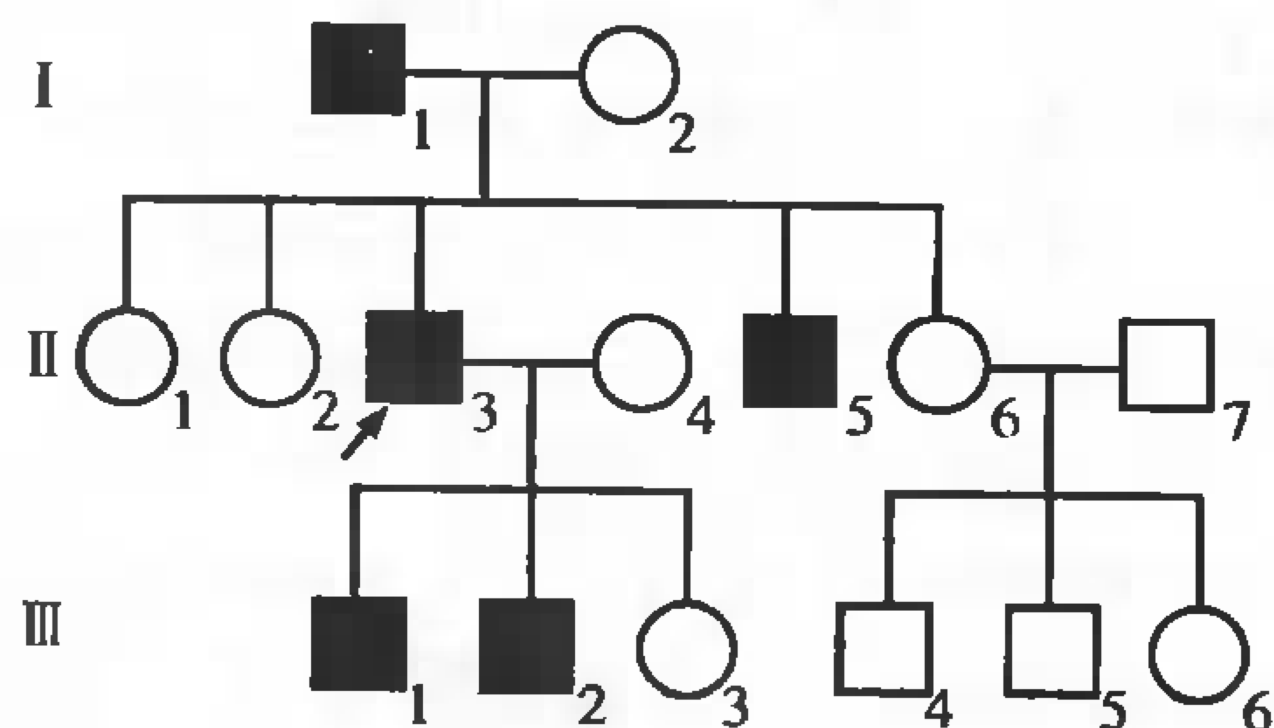


图4-12 一例脊髓小脑性共济失调的系谱

研究表明，遗传早现与有关基因的STR序列(尤其是三核苷酸重复)的动态突变有关。

六、限性遗传与从性遗传

限性遗传(sex-limited inheritance)是指位于常染色体上的某些基因的表达受性别的限制，只能在一种性别表达，在另一性别完全不表达。这与男女两性生理结构或性激素的差异有关。例如，子宫阴道积水、尿道下裂均为常染色体遗传病，致病基因虽在男女两性中都存在，但由于解剖结构不同，前者只见于女性，后者只见于男性。但不论男性还是女性，致病基因仍按常染色体遗传的规律传给子代。

从性遗传(sex-influenced inheritance)是指位于常染色体上的基因在不同性别有不同的表达程度和表达方式，从而造成男女性状分布上的差异。例如，血色素沉着症为AD病，人群中男性发病率高于女性。女性由于月经、流产或妊娠等生理或病理失血导致体内铁质减少，不易表现出症状，故该病多见于男性。

七、X染色体失活

根据Lyon假说，在胚胎发育早期，女性2条X染色体中的1条随机失活，因此，女性的2条X染色体存在嵌合现象。一般说来，女性一半细胞父源X染色体失活，另一半细胞母源X染色体失活。例如，一妇女为X连锁致病基因杂合子，预期半数细胞中带有突变基因的X染色体失活，细胞是正常的；另外半数细胞中带有正常基因的X染色体失活，细胞将为突变型。曾有报道，甲型血友病、Duchenne型肌营养不良(假性肥大性肌营养不良症的一种主要类型)男患者的杂合子母亲患同样的疾病。这是因为女性X染色体有随机失活现象，机遇使她的带有正常显性基因的那条X染色体在大部分细胞中失活，而带有隐性致病基因的那条X染色恰好有活性，从而表现出或轻或重的临床症状。

小　结

单基因遗传病是指由一对等位基因异常所引起的疾病。包括:常染色体显性遗传病(AD病)、常染色体隐性遗传病(AR病)、X连锁显性遗传病(XD病)、X连锁隐性遗传病(XR病)、Y连锁遗传病及线粒体遗传病。

AD病患者一般为杂合体,患者双亲之一发病,患者的同胞、子女1/2发病,男女发病机会均等,具有连续遗传现象。根据致病基因的表达差异,AD病又分为完全显性、不完全显性、共显性、不规则显性以及延迟显性等类型。

AR病患者双亲是肯定携带者,子女一般不发病;同胞1/4发病,但生育子女少的家庭,由于选择偏倚,发病风险往往偏高,无连续传递现象,男女发病机会均等;近亲婚配使子女发病风险增高,越是罕见的AR病,近亲婚配的危害性越大。

群体中XD病女患者是男患者的2倍。女患者与正常男性婚配,儿女发病的概率均为1/2;男患者与正常女性婚配,女儿全部患病,儿子都正常;存在连续传递和交叉遗传现象。

群体中XR病男患者多于女患者。男患者与正常女性婚配,女儿全部为携带者,儿子都正常;男患者与女性携带者婚配,儿女发病的概率均为1/2;女携带者与正常男性婚配,儿子1/2发病,女儿1/2为携带者。存在交叉遗传现象。

Y连锁遗传病只在男性个体中垂直传递,也称为全男性遗传或限雄性遗传。

基因组中决定某一性状或疾病表现的一对主基因外的其他基因称为这对基因的遗传背景或基因环境。基因多效性是指一个突变基因或一对基因可以产生多种表型效应。遗传异质性是指表型相似而基因型不同的现象,分等位基因异质性与位点异质性2种类型。基因组印记是指来自父方和母方的同源染色体或等位基因存在功能上的差异,传给子女时可形成不同表型的现象。遗传早现是指某些遗传病,在世代传递过程中,发病年龄逐代提前、病情逐代加重的现象。限性遗传是指位于常染色体上的某些基因的表达受性别的限制,只能在一种性别表达。从性遗传是指常染色体上基因在不同性别有不同的表达程度和表达方式,从而造成男女性状分布上的差异。

(梁素华)

第五章　线粒体遗传病

除哺乳动物成熟红细胞外，线粒体普遍存在于需氧呼吸的真核细胞的细胞质中，它是细胞物质氧化的主要场所和能量供给中心。线粒体内存在自身的遗传系统，是动物细胞核外惟一含 DNA 的细胞器。人类细胞的线粒体 DNA(mitochondrial DNA，mtDNA)是人类基因组的组成部分，被称为“25 号染色体”。线粒体基因涉及编码多种 tRNA、rRNA 及一些功能蛋白质，这些均参与维持线粒体的功能。早期有学者注意到某些疾病可能属细胞质遗传，1987 年 Wallace 明确提出线粒体 DNA 突变可引起人类疾病。最近十多年来，线粒体遗传及线粒体遗传病的研究发展十分迅速，现已确认了 100 多种致病点突变和 200 多种缺失、插入和重排，其中约 60%影响线粒体 tRNA，35%影响呼吸链多肽链亚单位，5%累及线粒体 rRNA。

第一节　线粒体基因组

一、线粒体基因组的结构

mtDNA 构成线粒体基因组。人的 mtDNA 是由 2 条 DNA 单链组成的环状闭合分子，外环为重链(H 链)，内环为轻链(L 链)，mtDNA 分子全长 16 569 bp。线粒体基因组含 37 个基因，编码 13 种多肽链、22 种 tRNA 和 2 种 rRNA。这 13 种多肽链都是呼吸链酶复合体的亚单位。mtDNA 2 条链的碱基组成差别很大，H 链富含 G，而 L 链多含 C。H 链编码 12 种多肽链、12 S rRNA、16 S rRNA 和 14 种 tRNA，而 L 链仅编码 1 种多肽链和 8 种 tRNA。mtDNA 没有内含子，惟一的非编码区是约 1 000 bp 的 D 环区，该区包含 mtDNA 重链复制的起始点、轻链转录的启动子以及 4 个高度保守序列和终止区。mtDNA 具有 2 个复制起始点，分别起始复制 H 链、L 链。它的转录则是由位于 D 环区的 2 个启动子同时开始的(图 5-1)。

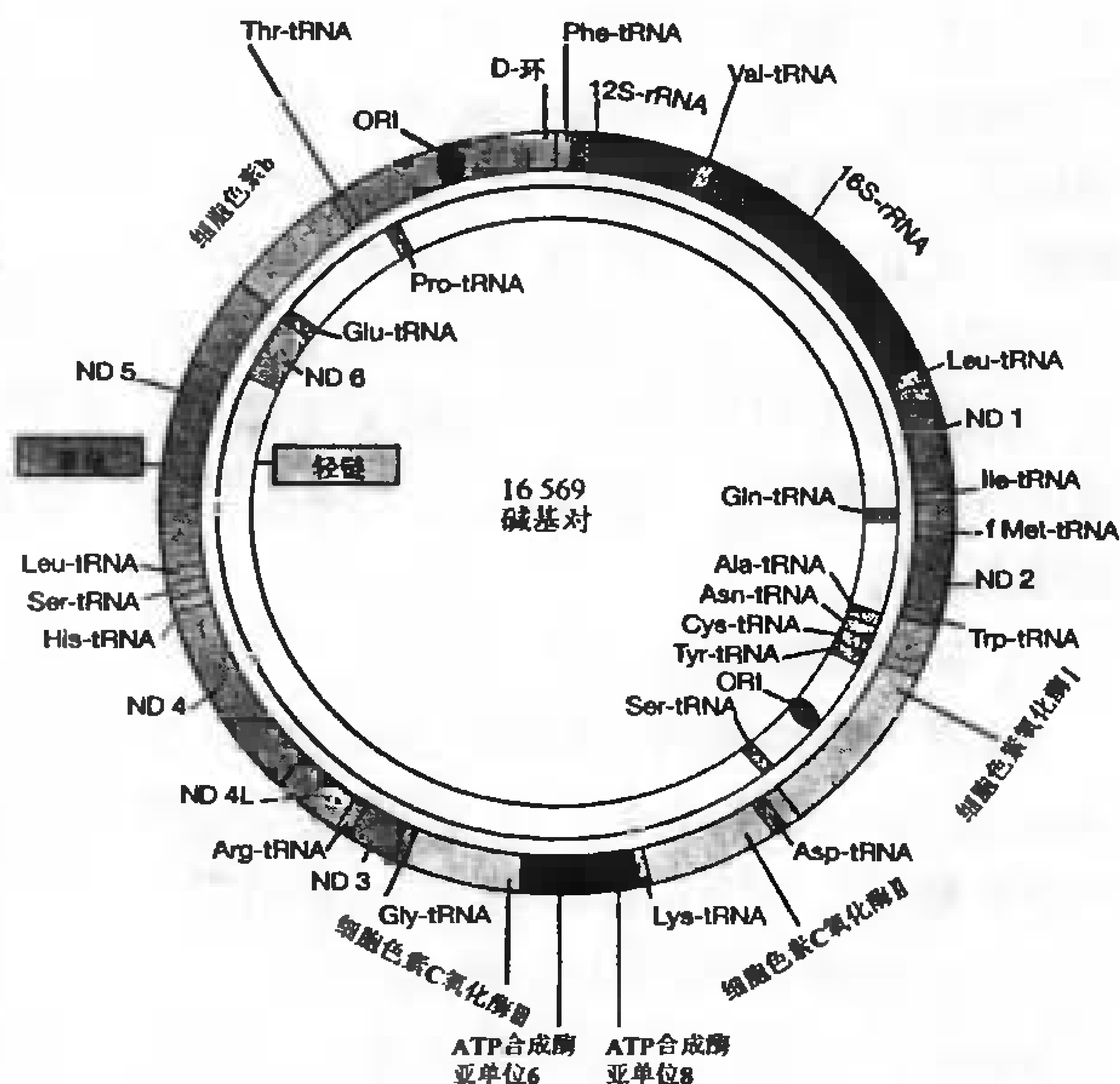


图 5-1　人类 mtDNA 的结构

mtDNA 分子上无核酸结合蛋白，缺少组蛋白的保护，而且线粒体中无 DNA 损伤修复系统，这就使 mtDNA

易于发生突变。所有线粒体均含多拷贝的 mtDNA，一个细胞内通常有数百个线粒体，每个线粒体内含 2～10 个 mtDNA 分子，因此每个细胞有数千个 mtDNA，而每个分子都可能发生突变，故 mtDNA 的突变率相当高。

二、线粒体基因组的遗传特点

与核 DNA 相比较，mtDNA 具有以下主要特点。

（一）半自主性

mtDNA 能够独立地复制、转录和翻译，但维持线粒体结构和功能的主要大分子复合物和大多数氧化磷酸化酶亚单位是由核 DNA 编码的，故其功能又受核基因的影响。

（二）遗传密码与通用密码不同

在线粒体遗传密码中，有 4 个密码子的含义与核基因的通用密码不同（表 5－1）。例如，UGA 编码色氨酸，而非终止密码。另外，线粒体的 tRNA 兼用性较强，仅用 22 个 tRNA 来识别 48 个密码子，而识别通用密码子的 tRNA 要多得多。

表 5－1　通用密码和线粒体密码的差异

密码子	通用密码的含义	线粒体密码的含义
UGA	终止	色氨酸
AUA	异亮氨酸	甲硫氨酸
AGA	精氨酸	终止
AGG	精氨酸	终止

（三）母系遗传

母亲将她的 mtDNA 传给她的所有子女，而只有她的女儿又将她的 mtDNA 传给下一代。这种传递方式称为母系遗传（maternal inheritance）。这是因为精卵结合时，精子提供的主要是核 DNA，受精卵的胞质绝大部分来自卵子，即受精卵中的 mtDNA 几乎都是母亲提供的。细胞分裂过程中，核的分裂是均等的，而细胞质的线粒体则是随机进入不同的子细胞。正是由于受精过程与细胞分裂期间线粒体与核的不同行为，以及每个线粒体含多个 mtDNA 拷贝，导致了线粒体遗传病的传递模式与经典的孟德尔遗传模式不同。实际上，目前发现的大部分 mtDNA 点突变都为母系遗传。因此，如果在某个家系中发现一些成员具有相同的临床表现，而且又是从受累的女性传递下来，就应考虑可能是由于 mtDNA 突变造成的。

（四）同质性与异质性

如果一个细胞或一种组织的所有 mtDNA 分子上的某一基因都是相同的，称为同质性（homoplasmy），或均为野生型，或均为突变型。如果 mtDNA 突变造成同一细胞或同一组织中野生型和突变型 mtDNA 共存，这种情况称为异质性（heteroplasmy）。异质性细胞经过有丝分裂和减数分裂，随机分配到 2 个子细胞中的突变型和野生型的 mtDNA 比例发生改变，mtDNA 基因型分别向纯合突变型和纯合野生型漂变，经过无数次分裂后细胞可达到同质性。氧化磷酸化缺陷程度与突变型 mtDNA 的比例成正比，氧化磷酸化酶活性的变化范围可在正常活性的 0～100%。异质性和有丝分裂、减数分裂过程的复制分离使相同核基

因型的细胞或个体,如同卵双生子,可具有不同的细胞质基因型,从而具有不同的表型。

(五)阈值效应

mtDNA 突变所致异常表型的出现,是由某种组织野生型与突变型 mtDNA 的相对比例以及该组织对能量的依赖程度决定的。突变的 mtDNA 数量达到一定程度时,才引起某种组织或器官的功能异常,这称为阈值效应。不同组织和器官对能量的依赖程度是不同的,脑、骨骼肌、心、肾、肝对能量的依赖性依次降低。当 mtDNA 突变,ATP 产生减少,低于维持各组织、器官正常功能最低值时,临床就有异常表现。最先受损的是中枢神经系统,其后为肌肉、心、肾和肝等。ATP 产生越少,累及器官越多,症状越严重。然而,阈值的大小还受其他多种因素的影响。不同的 mtDNA 突变基因阈值的大小不同,应用胞质杂种细胞系进行的研究发现,tRNA 基因点突变的阈值为 90%,而 mtDNA 大片段缺失突变的阈值为 60%。细胞核的遗传背景也与阈值的大小有关,对一些突变的表型而言,要求 100%的突变率,但有时即使达到 100%的突变率也不一定产生疾病表型,可见突变 mtDNA 的比率虽然是阈值效应的重要因素,但不是惟一因素。

(六)突变率高

mtDNA 的突变率比核 DNA 高 10~100 倍。mtDNA 的高突变率造成个体及群体中 mtDNA 序列的极大不同,任何两个人的 mtDNA,平均每 1 000 个碱基对中有 4 个不同。人群中存在多种无害到中度有害的 mtDNA 突变,且高度有害的 mtDNA 突变不断增多。但有害的突变会通过选择而消除,故线粒体遗传病表型并不常见,突变的基因型却很普遍。

第二节 线粒体基因突变与疾病

人类 mtDNA 突变与疾病间的关系是近年来分子病理学研究的热点之一。鉴于线粒体在能量代谢中的重要作用,mtDNA 的高频率突变,无内含子及无完整的修复系统等特征,mtDNA 突变在许多疾病中存在,这些疾病包括具有母系遗传特征的疾病,中老年人发生的一些退行性疾病,甚至衰老过程等。mtDNA 突变所表现出的一些临床特征包括:肌病、心肌病、痴呆、突发性肌阵挛、耳聋、失明、贫血、糖尿病和大脑供血异常等。这些临床缺陷的形成与严重程度依赖于多种因素。例如,胚胎发育早期线粒体突变基因组的复制分离程度,突变的线粒体基因在某一特定组织中存在的数量,以及在临床上出现异常之前组织中突变的线粒体 DNA 所需达到的阈值水平等。另外,值得注意的是,氧化磷酸化过程中 5 种酶复合物是 mtDNA 和核 DMA 共同编码的,相应核基因突变也可产生类似于线粒体疾病的表现,而有些线粒体疾病是 mtDNA 突变和核基因突变共同作用的结果。

一、线粒体基因突变的类型

(一)错义突变

错义突变又称氨基酸替代突变,是指 mtDNA 分子中编码 13 种多肽链的基因点突变。这些突变主要与脑、脊髓及神经性疾病有关,如 Leber 遗传性视神经病、神经肌病等。

(二)蛋白质生物合成基因突变

蛋白质生物合成基因突变是指 mtDNA 分子中的 tRNA 基因突变。突变的结果是：mtDNA 的结构改变直接影响 tRNA 携带氨基酸的功能，间接影响线粒体蛋白质的合成。这类突变所致的疾病比错义突变所致的疾病更具有系统性特征。主要的疾病有肌阵挛性癫痫伴碎红纤维病、线粒体肌病脑病伴乳酸中毒及中风样发作综合征、母系遗传的心肌病等。

(三)缺失、插入突变

以缺失突变相对多见。缺失通常是由于 mtDNA 的异常重组或在复制过程中异常滑动所致。这类突变存在于许多神经肌肉性疾病及一些退化性疾病、肾病和肝病中，甚至衰老也与之有关。缺失突变导致的疾病一般无家族史。发生在胚胎早期较大片段的 mtDNA 缺失将导致胚胎死亡；发生在胚胎发育晚期的 mtDNA 缺失受到正常 mtDNA 的功能弥补，最初不表现出症状或表现出极轻症状，但病情往往随年龄增长而逐步加重，这主要是由于缺失的 mtDNA 渐进式增殖，其比例逐渐增加所致。

(四)mtDNA 拷贝数目突变

这是指 mtDNA 拷贝数大大低于正常。这种突变较少，仅见于一些致死性婴儿呼吸障碍、乳酸中毒或肌病、肝肾功能衰竭的病例。

二、常见线粒体遗传病

(一)Leber 遗传性视神经病

Leber 遗传性视神经病(Leber hereditary optic neuropathy，LHON)是一种罕见的眼部线粒体疾病，是人类母系遗传的典型病例。本病起病为急性或亚急性眼球后神经炎，导致双侧视神经萎缩，首发症状多为视物模糊，接着在几个月内出现无痛性、完全或接近完全的失明，两眼常同时受累，或一眼失明不久另一眼也很快失明。其他伴随症状包括反射亢进、小脑共济失调、心律紊乱等。发病高峰年龄是 18～30 岁，但任何年龄均可发病。男女发病比例为 4∶1。

LHON 是在人类发现的第 1 种母系遗传病，迄今尚未发现有男性患者将此病传给后代的例外情况。现已发现 mtDNA 上许多位点的突变与 LHON 有关。1987 年，Wallace 最先发现该病 mtDNA 11 778 位点的 G→A 突变，突变使 NADH 脱氢酶亚单位 4(ND_4)中第 340 位精氨酸变成了组氨酸，从而影响了线粒体能量的产生。大约 50% 的 LHON 由该位点突变引起。除此之外，现在已发现 10 多个点突变存在于该病中，它们分布于 ND_1、ND_2、ND_3、ND_4、ND_5 和细胞色素 b 等基因位点内，而且这些突变都有累积作用。这表明 LHON 患者视力的丧失是电子传递链功能受阻，而非某种特异性酶缺陷。由于电子传递链的多肽链是由 mtDNA 和核 DNA 共同编码的，故核基因也可能影响 LHON 的发生。

利用 11 778 位点点突变可以对一半的 LHON 家系进行基因诊断。其原理为：由于 11 778位点 G→A 转换，丧失了 SfaN Ⅰ 酶切位点，同时又获得了 MaeⅢ酶切位点，因此，可以在11 778位点两侧设计一对引物进行 PCR 扩增，其产物用 SfaN Ⅰ 酶切，正常个体的 PCR 产物可被切成 2 条带，而患者和携带者只有 1 条带。如果用 MaeⅢ酶切 PCR 产物，则正好

相反，正常个体呈现 1 条带，患者呈现 2 条带，携带者呈现 3 条带。

（二）肌阵挛性癫痫伴碎红纤维病

肌阵挛性癫痫伴碎红纤维病（myoclonus epilepsy and ragged red fibers，MERRF）是一种线粒体肌病。患者具有多系统紊乱的表现，包括肌阵挛性癫痫的短暂发作、不协调肌肉运动（共济失调）、肌细胞减少（肌病）、轻度痴呆、耳聋和脊髓神经的退化等。骨骼肌活检可见破碎红纤维，这是由于大量的异常线粒体堆积在肌细胞中，特异性染料将其染成红色所致。本病发病年龄一般为童年，病情可持续若干年，有明显的母系遗传特点，患者的母系亲属常有脑电图异常、感觉神经听力丧失、痴呆、呼吸异常、扩张性心肌病等临床表现。

80%～90%的 MERRF 患者 mtDNA 的 $tRNA^{Lys}$基因存在 8 344 位点 A→G 点突变，少数患者是该基因 8 356 位点 T→C 点突变。这些突变使 tRNA 结构改变，影响了线粒体蛋白的整体合成水平，产生了一系列 MERRF 特定的翻译产物，除酶复合物Ⅱ外，所有氧化磷酸化成分含量都减少。

MERRF 家族成员 mtDNA 通常为异质性，氧化磷酸化酶水平随年龄的增大迅速降低。通常 20 岁以下的 MERRF 患者，85%的 mtDNA 突变时表现仍正常，mtDNA 突变达 95%时才会出现 MERRF 临床表现；60 岁以上的患者 mtDNA 突变在 63%时就表现中度异常，85%的 mtDNA 突变时则有相当重的临床表现。

（三）线粒体肌病脑病伴乳酸中毒及中风样发作综合征

线粒体肌病脑病伴乳酸中毒及中风样发作综合征（mitochondrial encephalomyopathy with lactic acidosis and stroke－like episodes，MELAS）患者常在 40 岁以前出现异常，表现为突发呕吐、复发性休克、肌肉组织病变等。少数患者伴痴呆、耳聋、周围性偏头痛、眼外肌无力或麻痹和身体矮小等。由于线粒体的功能障碍影响丙酮酸的代谢，大量丙酮酸生成乳酸并积聚在血液和体液中，导致乳酸性酸中毒。MELAS 患者的特征性病理变化是在脑和肌肉的小动脉和毛细血管管壁中有大量形态异常的线粒体集聚。

约 80%MELAS 患者 mtDNA 编码的 *tRNA* 基因 3 243 位点存在 A→G 点突变，少见的点突变发生在该基因的 3 291、3 271、3 256 和 3 252 位点及 $tRNA^{Val}$基因、细胞色素氧化酶Ⅲ基因上。发生在 $tRNA^{leu(UUR)}$基因的点突变改变了 $tRNA^{leu(UUR)}$的结构，并使该 *tRNA* 基因和 *rRNA* 基因下游紧密结合的转录终止子失活，因此 MELAS 突变可降低 mtDNA 的转录活性并改变线粒体 rRNA 和 mRNA 转录的比例。

不同种类的线粒体突变所导致的临床变异是复杂的。在一些 3 243 位点 A→G 突变的个体中，惟一的表型特征是糖尿病和耳聋，而在 3 250、3 251、3 302、3 303 和 3 260 位点突变的患者中，肌病是主要特征。3 256 位点 C→T 突变的患者则表现出 MELAS 和 MERRF 2 种疾病的表现。总之，不同的线粒体 *tRNA* 基因突变可引起不同的功能紊乱，一些线粒体 *tRNA* 基因突变能产生相似的临床表现，而同一 *tRNA* 基因不同位点的突变也可产生不同的临床表现。

（四）慢性进行性眼外肌麻痹

慢性进行性眼外肌麻痹又称 Kearns－Sayre 综合征（Kearns－Sayre syndrome，KSS）。常见临床表现是进行性眼外肌麻痹和视网膜色素变性。KSS 的表现还包括心肌传导功能异常、共济失调、耳聋、痴呆和糖尿病。发病年龄一般低于 20 岁，大多数患者在确诊后几年

内死亡。

KSS并不表现出特定的母系遗传或核基因遗传方式，然而其临床表现表明它是一种线粒体遗传病。mtDNA分析表明：KSS患者存在mtDNA结构改变，主要是大片段的缺失。缺失大多发生在重链与轻链的2个复制起始点之间。最常见的是8 468位点和13 446位点之间的4 977 bp的缺失，该缺失的断裂点分别位于ATP酶亚单位8基因和ND_5基因内，该缺失占KSS的1/3。最大片段的缺失为5 786位点和15 944位点之间大于10 kb的缺失。由于缺失，丢掉了许多基因，尤其是*tRNA*基因，因此患者都有不同程度的mtDNA蛋白合成缺陷。

KSS患者病情严重程度与异质性程度、发生改变的mtDNA的组织分布有关。当肌细胞中缺失突变的mtDNA达到85%时，可呈现所有KSS的临床特征；在异质性处于较低水平时，进行性眼外肌麻痹是主要表现；当造血干细胞中存在大量的mtDNA缺失突变时，就会表现出一种致命且早发的疾病，称Pearson综合征。该综合征的主要特点是血细胞不能利用铁合成血红蛋白，从而引起缺铁性贫血。

（五）糖尿病

随着对线粒体遗传病的认识不断深入，线粒体基因突变被认为是糖尿病的一种新的遗传缺陷。线粒体基因突变导致的糖尿病是一种少见亚型，美国糖尿病学会1997年糖尿病诊断与分型标准将其作为一种新的亚型，正式列入“其他特殊类型糖尿病”的分型中，属β细胞遗传缺陷性疾病。

近年来，已发现与糖尿病相关的20余种线粒体基因突变，除插入、缺失、重排等突变方式可能引起胰岛素β细胞缺陷外，常见的突变方式为线粒体tRNA基因的点突变，尤其是携带亮氨酸、丝氨酸、谷氨酸及赖氨酸的tRNA的基因突变率较高。其中最常见的是tRNA基因3 243位点A→G突变，该突变降低了mtDNA编码蛋白的稳定性和相应氨基酸与其tRNA的结合力，从而影响了线粒体蛋白质的合成。众多研究发现，伴3 243位点A→G突变的糖尿病患者呈不同程度的异质性，突变的mtDNA在32%～63%不等。自1992年Vanden Duwaland等报道1个tRNA基因3 243位点A→G突变所致糖尿病伴耳聋的家系以来，3 243位点的突变一直是各国学者研究的热点。综合文献分析，全球糖尿病人群中，线粒体基因突变引起者约为1.5%，其中线粒体tRNA基因3 243位点突变引起的糖尿病占mtDNA突变所致糖尿病的50%左右。

近年的流行病学研究还证实，2%～4%的糖尿病患者带有ND_1基因3 316位点G→C突变。该突变使ND_1亚单位上的一个疏水性亮氨酸变成亲水性苏氨酸，改变了ND_1的空间结构，使呼吸链酶活性降低，导致ATP合成障碍。

三、线粒体基因突变与衰老

20世纪80年代以来，许多学者对mtDNA突变和衰老关系进行了研究。例如，有人应用PCR方法检测正常成人的心肌和脑组织的mtDNA突变，发现mtDNA有少量特异性缺失，而胎儿的心肌和脑组织中没有发现这类缺失，提示随年龄的增长可能伴有mtDNA损伤。还有实验证明，老年大鼠的mtDNA积累了各种重排。因此，人们推测衰老可能与异常mtDNA积累有关。此外，在mtDNA缺失所致疾病中，大多数在成年期才表现出症状，且

随年龄增长症状逐渐加重。这可能是由于 mtDNA 突变的积累，使线粒体氧化磷酸化的能力逐渐降低，细胞产生 ATP 越来越少之故。

小　结

mtDNA 是由 2 条链构成的闭合环状分子，长度为 16 569 bp，含 37 个基因，编码 13 种呼吸链酶复合体亚单位、22 种 tRNA 和 2 种 rRNA。

mtDNA 分子上无核酸结合蛋白，缺少组蛋白的保护，而且无 DNA 损伤修复系统。

mtDNA 与核基因比较在遗传上有以下特点：①功能上具半自主性，能够独立复制和表达，同时又受核基因的影响；②一些遗传密码与通用密码不同；③表现为母系遗传；④在同一组织或个体中存在同质性或异质性；⑤有阈值效应，突变的 mtDNA 数量达到不能维持组织器官功能的最低能量需要时，就会出现受损的表型；⑥突变率比核基因高 10～100 倍。

线粒体基因突变有 4 类，即：错义突变，蛋白质合成基因突变，插入、缺失突变和 mtDNA拷贝数目突变。

mtDNA 分子病理学研究证实，在许多疾病中都存在线粒体遗传结构改变。这些疾病包括具有母系遗传特征的疾病，中老年发生的退行性疾病等。mtDNA 突变表现出的临床特征有肌病、心肌病、失明、贫血、突发性阵挛等。

常见的线粒体遗传病有：①Leber 遗传性视神经病，该病的临床表现与 mtDNA 许多位点的错义突变有关，最多见的是 11 778 位点的碱基替代；②MERRF，该病有典型母系遗传特征，80%～90%患者 mtDNA 的 *tRNA* 基因 8 344 位点存在 A→G 突变；③MELAS，有脑病、肌肉组织病变、乳酸中毒等临床特征，80%患者 *tRNA* 基因 3 243 位点存在 A→G 突变。④KSS，常见临床表现是进行性眼外肌麻痹和视网膜色素变性，该病的 *mtDNA* 改变主要是大片段的缺失。最常见的缺失是 8 468 位点和 13 446 位点之间的 4 977 bp 的缺失。此外，糖尿病的一些类型中已发现 20 余种线粒体基因突变；衰老过程中亦伴随 *mtDNA* 的缺失突变。

（税青林）

第六章　多基因遗传病

人类单基因遗传病的发生是一对基因作用的结果。虽然在某些情况下，修饰基因(modifier)或环境因素对表型也有一定的影响，但总的说来，单基因遗传病的发生还是一对基因(即主基因)起主要作用。人类另一些有遗传基础的常见病，如原发性高血压、糖尿病、精神分裂症等，是受多对等位基因控制的，并且环境因素起着较为重要的作用。所以，这类疾病称为多基因遗传病。近年来研究表明，多基因遗传病除受多对基因和环境因素作用外，也存在主基因(又称易感主基因)的作用。因此，多基因遗传病又称为复杂疾病(complex disease)。

第一节　数量性状的遗传

一、质量性状与数量性状

单基因遗传的性状(或疾病)主要与一对主基因的作用有关，相对性状差异显著，如豌豆的圆与皱，疾病的有与无。因此，单基因遗传的性状称为质量性状(qualitative character)。在完全显性或隐性情况下，性状在群体中的变异表现为2个峰：AA和Aa(3/4)，aa(1/4)；在不完全显性情况下表现为3个峰：AA(1/4)，Aa(1/2)和aa(1/4)；峰与峰之间存在质的差异，即群体中的变异分布是不连续的(图6-1A，B)。

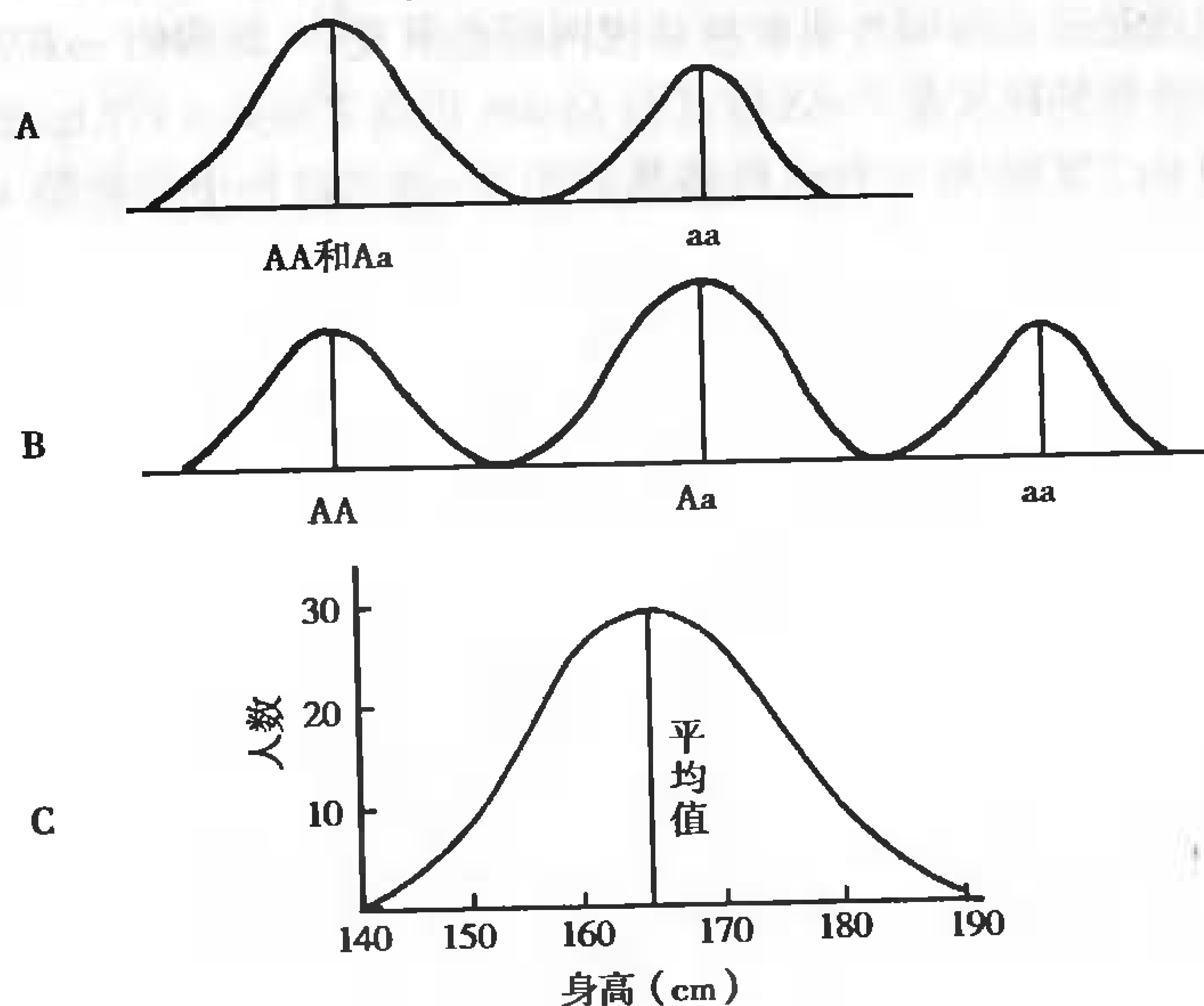


图6-1　质量性状与数量性状的变异

A、B：质量性状　C：数量性状

多基因遗传(polygenic inheritance)的性状是多对基因和环境因素相互作用的结果,性状在群体中的变异呈现连续的数量级差,无质的不同,表现为一个峰(图 6 - 1C),故多基因遗传的性状称为数量性状(quantitative character)。

二、多基因假说

多基因遗传的现象是瑞典遗传学家 Nilsson - Ehle 在 1909 年研究白色和暗红色小麦的杂交实验时发现的,据此提出了多基因假说。其要点是:①数量性状的遗传物质也是基因,但不是 1 对基因,而是 2 对以上的基因;②这些基因彼此之间没有显隐性之分,是共显性的;③每对等位基因中的单个基因对表型的作用是微小的,称其为微效基因(minor gene);但多对微效基因的作用累加起来,可形成明显的表型效应,称累加效应(additive effect),故微效基因也称加性基因(additive gene);④表型效应除受多对微效基因的作用影响外,环境因素也起作用。

三、数量性状的遗传

(一)多基因遗传的特点

多基因遗传,即数量性状的遗传,除受多对等位基因作用外,还受环境因素的影响。其特点是:2 个极端变异的个体杂交,子 1 代都是中间类型,但由于受环境因素的影响,子 1 代也有一定范围的变异;2 个中间类型的子 1 代个体杂交后,子 2 代大部分仍是中间类型,但变异范围要比子 1 代广泛,可产生出近于极端变异的个体,除受环境因素的影响外,基因的分离和自由组合对变异的产生也有作用;在随机杂交的群体中,变异范围广泛,大多数个体近于中间类型,极端变异的个体很少,遗传因素和环境因素对变异的产生都有影响。

(二)数量性状的遗传

人的身高、血压、智力、肤色等都属于数量性状。现以人类身高的遗传为例说明多基因遗传的特点。人的身高受控于多对微效基因,假定受控于 3 对非连锁的微效基因 AA′、BB′和 CC′,每个基因的频率均为 0.50,其中基因 A、B、C 较基因 A′、B′、C′对身体长高有增强作用。将基因 A、B、C 称为高基因,基因 A′、B′、C′称为矮基因。在不考虑环境因素影响的情况下,基因型 AA′BB′CC′的个体表现为中等身高(165 cm)。在此基础上,假定每增加 1 个高基因或减少 1 个矮基因,将导致身高 5 cm 的改变。那么,基因型 AABBCC 和基因型 A′A′B′B′C′C′的个体,身高分别为 195 cm、135 cm,为群体中的极高个体和极矮个体。如果这 2 种基因型的个体婚配,子 1 代基因型均为 AA′BB′CC′,理论上,子 1 代均为中等身高者,但由于受环境因素的影响,子 1 代身高仍会有差异。如果子 1 代个体间婚配,子 2 代的大部分个体仍具有中等身高,但变异范围广泛,将会出现某些极高个体和极矮个体。这种变异一是受 3 对基因分离和自由组合的影响,二是环境因素的作用。

就 3 对基因分离和自由组合而言:子 1 代可产生 8 种配子;子 1 代间婚配时,子 2 代形成 64 种基因组合,包含 27 种基因型;在 27 种基因型中,根据所含高基因(或矮基因)的数目,可归并为 7 种表型;从矮到高的表型依次为:0,6′(0 个高基因,6 个矮基因)、1,5′、2,4′、3,3′、4,2′、5,1′、6,0′,其频数分别为 1、6、15、20、15、6、1(表 6 - 1)。依据子 2 代各表型

的频数绘制柱形图,可发现近似正态分布(图 6-2)。如果影响身高的基因对数越多,则曲线越接近正态分布。

表 6-1　子 2 代身高基因的组合

亲代配子	亲代配子							
	ABC	A′BC	AB′C	ABC′	A′B′C	AB′C′	A′BC′	A′B′C′
ABC	AABBCC	A′ABBCC	AAB′BCC	AABBC′C	A′AB′BCC	AAB′BC′C	A′ABBC′C	A′AB′BC′C
A′BC	A′ABBCC	A′A′BBCC	A′AB′BCC	A′ABBC′C	A′A′B′BCC	A′AB′BC′C	A′A′BBC′C	A′A′B′BC′C
AB′C	AAB′BCC	A′AB′BCC	AAB′B′CC	AAB′BC′C	A′AB′B′CC	AAB′B′C′C	A′AB′BC′C	A′AB′B′C′C
ABC′	AABBC′C	A′ABBC′C	AAB′BC′C	AABBC′C′	A′AB′BC′C	AAB′BC′C′	A′ABBC′C′	A′AB′BC′C′
A′B′C	A′AB′BCC	A′A′B′BCC	A′AB′B′CC	A′AB′BC′C	A′A′B′B′CC	A′AB′B′C′C	A′A′B′BC′C	A′A′B′B′C′C
AB′C′	AAB′BC′C	A′AB′BC′C	AAB′B′C′C	AAB′BC′C′	A′AB′B′C′C	AAB′B′C′C′	A′AB′BC′C′	A′AB′B′C′C′
A′BC′	A′ABBC′C	A′A′BBC′C	A′AB′BC′C	A′ABBC′C′	A′A′B′BC′C	A′AB′BC′C′	A′A′BBC′C′	A′A′B′BC′C′
A′B′C′	A′AB′BC′C	A′A′B′BC′C	A′AB′B′C′C	A′AB′BC′C′	A′A′B′B′C′C	A′AB′B′C′C′	A′A′B′BC′C′	A′A′B′B′C′C′

人类虽不存在上述假设的婚配情况,但由于大多数人为不同程度的杂合基因型,且大多数人为中等身高,故随机婚配将类似于上述子 1 代个体间的婚配,也就是说,随机婚配繁衍的后代将与上述子 2 代的变异分布相似。

在此正态分布的曲线中,可见均数处曲线最高,以均数(中等身高)为对称轴,两侧曲线对称分布;越远离均数,变异的分布越少,曲线越低。这说明具有中等身高的个体最多,极高个体和极矮个体很少。统计学理论表明,在正态分布中,群体中变异分布与正态曲线下面积有关:将正态曲线总面积(100%)比作群体中变异的全部个体,那么,均数(μ)± 1 倍标准差(δ)包括正态曲线下面积的 68.28%,此范围以外的面积占 31.72%,左侧和右侧各占 15.86%;$\mu \pm 2\delta$ 包括总面积的 95.46%,此范围以外的面积占 4.54%,左侧和右侧各占 2.27%;$\mu \pm 3\delta$ 包括总面积的 99.74%,此范围以外的面积占 0.26%,左侧和右侧各占 0.13%(图 6-3)。

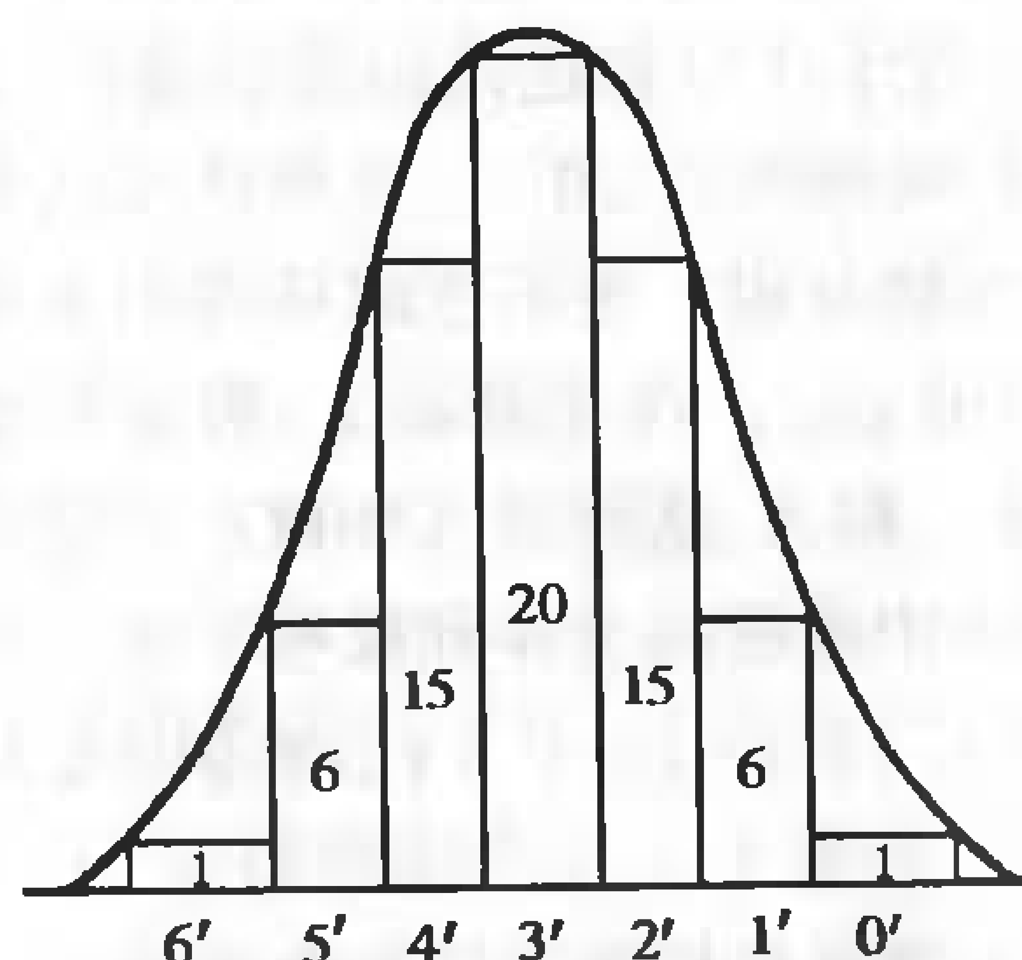


图 6-2　子 2 代身高变异分布

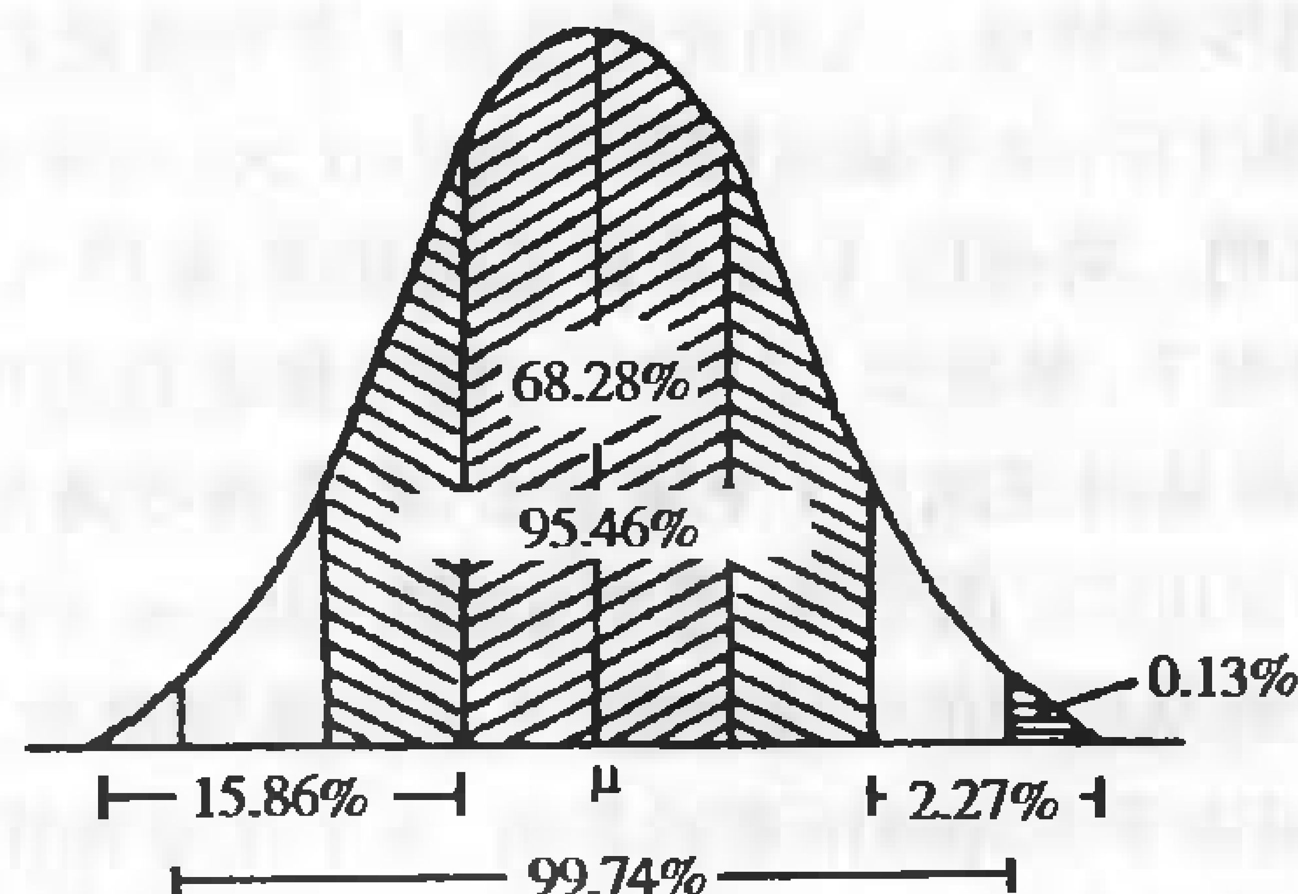


图 6-3　正态曲线下面积的分布规律

第二节　多基因遗传病

人类一些常见的疾病和先天性畸形群体发病率为0.1%～1%，患者一级亲属发病率可达1%～10%。这说明这类疾病和先天性畸形呈现家族聚集现象，有遗传基础，子代从患病亲代得到致病基因组合的可能性增高。现认为这类疾病和先天性畸形属于多基因遗传病。在多基因遗传病的发病中，不但遗传因素起作用，而且环境因素也起重要作用。

一、易感性、易患性和阈值

多基因遗传病的发生是遗传基础和环境因素共同决定的。其中由遗传基础决定的个体患病的风险称为易感性(susceptibility)，而一个个体在遗传基础和环境因素共同作用下患某种多基因遗传病的风险称易患性(liability)。通俗地说，易患性就是一个个体患某种多基因遗传病的可能性。当一个个体的易患性高达一定水平，即达到一定限度时，这个个体将发病，这个使个体发病的易患性限度称阈值(threshold)。群体中每个个体的易患性是不同的，群体的易患性变异也像多基因遗传的性状那样，呈正态分布，即大部分个体易患性接近平均值，易患性很高和很低的个体很少(图6-4)。很显然，阈值将连续分布的易患性变异分成两部分，即阈值及阈值以上的患者和阈值以下的健康者；即使同是健康者，易患性也会存在差异。患者与群体总人数的比率为群体发病率，也就是阈值右侧曲线下的面积与正态分布曲线下总面积(100%)的比率。在群体中，如果环境因素相同，阈值则代表着发病所需的最低的致病基因数。

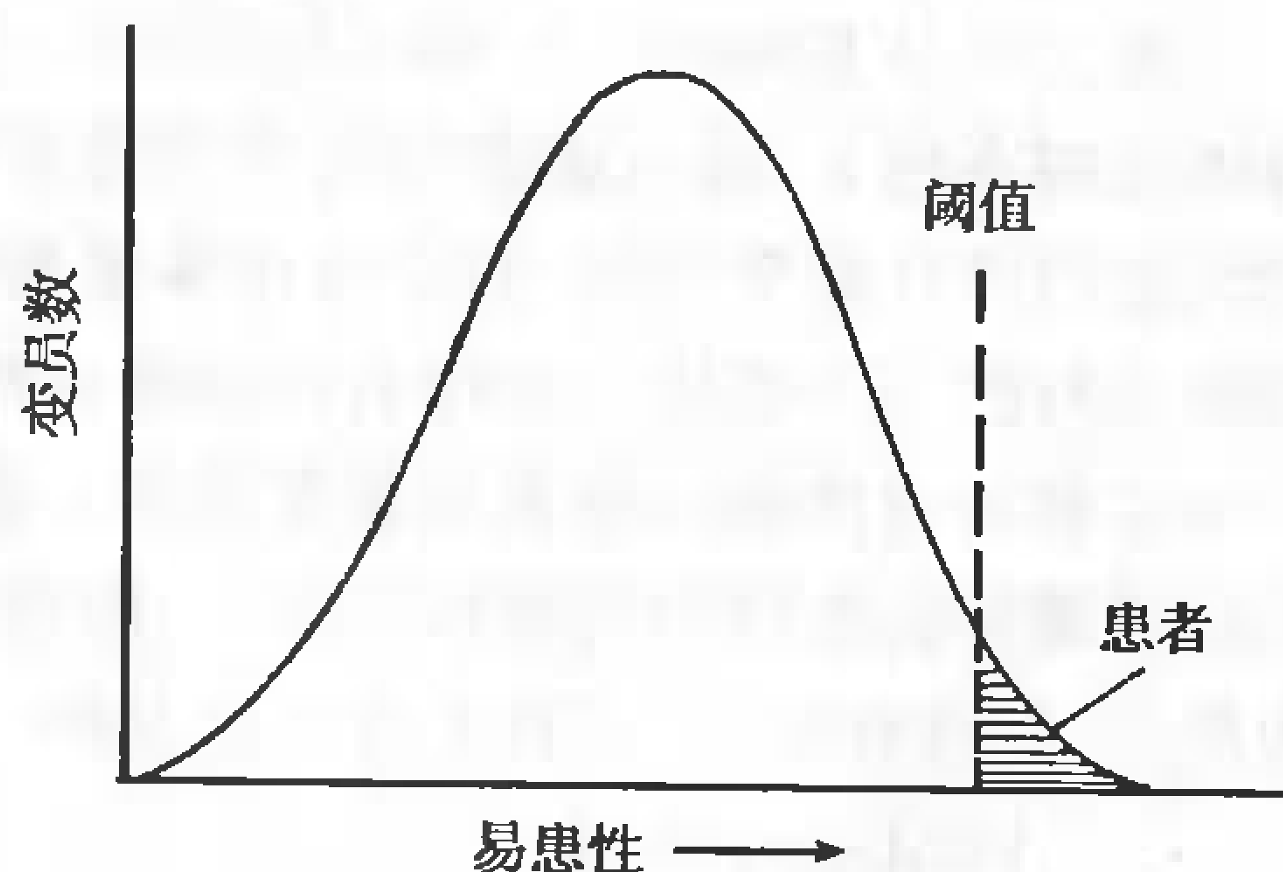


图6-4　群体的易患性变异

一个个体如果患某种多基因遗传病，说明他的易患性达到或超过阈值；如果未患病，则说明他的易患性未达到阈值。一个个体易患性究竟有多高，无法准确测量，只能根据其家庭成员的发病情况作粗略估计(后述)。然而，一个群体的易患性平均值则可以从该群体的发病率做出估计。根据正态曲线下面积分布规律，可由群体发病率估计阈值与易患性平均值之间的距离，二者之间的距离反映了群体易患性的高低。易患性变异的正态分布曲线下面积代表群体的总人数，阈值右侧曲线下的面积代表患者人数，该面积占总面积的百分比就是群体发病率。根据群体发病率，通过查“正态分布 X 和 a 值表”就可得出阈值距离易患性平均值有几个标准差。例如，精神分裂症的群体发病率为1.0%，经查表(表从略)可知，其阈值距易患性平均值为2.3个标准差，即距易患性平均值右侧2.3个标准差处是该病的阈值。因此：一种多基因遗传病群体发病率越低，说明其阈值越高，阈值与易患性平均值间距离也就越远，相对来说，可以认为该群体该病的易患性平均值低；相反，群体发病率越高，说明阈值越低，阈值与易患性平均值间距离也就越近，相对来说，群体的易患性平均值越高；假如群体发病率相当高，以至于阈值等于易患性平均值，二者之

间已不存在距离，表示群体中有一半的人是患者，说明该群体的易患性相当高。

二、遗传率

多基因遗传病受遗传基础和环境因素的共同作用，其中遗传基础在多基因遗传病的发病中所起作用的大小称为遗传率或遗传度(hebitability)，通常用百分率(%)表示。一种多基因遗传病的发病如果完全是遗传基础起作用，环境因素不起作用，遗传率就是100%，这种情况实际上是不存在的。一般来说，70%~80%的遗传率就表明遗传基础在发病上起重要作用，环境因素起的作用较小，认为此遗传率是比较高的；相反，遗传率为30%~40%表明遗传基础所起作用较小，环境因素在疾病的发生上起重要作用，认为此遗传率是比较低的。

遗传率有广义和狭义之分。广义遗传率是指全部遗传因素在多基因遗传病中所起作用的大小，用 H 表示；而狭义遗传率仅指微效基因所起作用的大小，用 h^2 表示。多基因遗传病的遗传率计算方法有2种。

(一)Falconer 公式

此公式是 Falconer 于1965年提出的，其基本原理是：由于遗传率与患者亲属发病率、群体发病率相关，故可根据患者亲属发病率、群体发病率计算遗传率。如果一种多基因遗传病的群体发病率很低，而患者亲属发病率很高，即二者差距很大，这说明患者亲属的发病很大程度上是遗传了亲代的致病基因所致，也即该多基因遗传病遗传率高；相反，如果一种多基因遗传病的患者亲属发病率与群体发病率相差不大，则表明亲代致病基因在该多基因遗传病发病中所起作用不大，即该多基因遗传病的遗传率较低。由于计算过程较为复杂，故计算公式、计算方法此处从略。

(二)Holgiger 公式

此公式是 Holgiger 于1929年建立的，其计算遗传率的原理是：一卵双生的两个人具有完全相同的遗传结构，如果某种多基因遗传病的遗传因素所起作用越大，当一卵双生中的一个人患该病时，另一个人也患该病的可能性(发病一致率)必然越大；二卵双生的两个人在遗传上相当于同胞，其发病一致率必定低于一卵双生的发病一致率。多基因遗传病的遗传率越高，一卵双生发病一致率(C_{MZ})与二卵双生发病一致率(C_{DZ})之差必然越大。Holgiger 公式是：

$$h^2 = \frac{C_{MZ} - C_{DZ}}{1 - C_{DZ}}$$

例如，对躁狂抑郁性精神病的调查表明，一卵双生的15对中10对发病，发病一致率(C_{MZ})67%，二卵双生的40对中2对发病，发病一致率(C_{DZ})5%，求遗传率。

将 $C_{MZ}=67\%$，$C_{DZ}=5\%$代入上式，得：

$$h^2 = \frac{C_{MZ} - C_{DZ}}{1 - C_{DZ}} = \frac{67\% - 5\%}{1 - 5\%} = 65\%$$

这说明，在多基因遗传病中，遗传因素所起作用较环境因素为大。

常见多基因遗传病和先天性畸形的遗传率见表6-2。

表 6-2　常见多基因遗传的疾病(或畸形)的遗传率和发病率

疾病名称	群体发病率(%)	先证者一级亲属发病率(%)	男:女	遗传率(%)
哮喘	4.0	20	0.8	80
精神分裂症	1.0	16	1	80
原发性高血压	4~8	20~30	1	62
消化性溃疡	4.0	8	1	37
强直性脊椎炎	0.2	7(男性先证者) 2(女性先证者)	0.2	70
Ⅰ型糖尿病	0.2	2~5	1	75
Ⅱ型糖尿病	2~3	10~15	1	35
冠心病	2.5	7	1.5	65
先天性幽门狭窄	0.3	2(男性先证者) 10(女性先证者)	5.0	75
先天性髋关节脱位	0.07	4	0.2	70
先天性畸形足	0.1	3	2.0	68
先天性巨结肠	0.02	2(男性先证者) 8(女性先证者)	4.0	80
无脑畸形	0.5	2	0.5	60
脊柱裂	0.3	4	0.8	60
唇裂、唇裂+腭裂	0.17	4	1.6	76
腭裂	0.04	2	0.7	76

三、多基因遗传病的遗传特点

多基因遗传病与单基因遗传病相比,有明显不同的遗传特点,主要表现在以下几方面。

(一)家族聚集现象

患者家族成员的发病率高于群体发病率(0.1%~1%),但不符合任何一种单基因遗传病的遗传方式,患者一级亲属发病率为1%~10%,远低于单基因遗传病患者同胞(一级亲属)1/2(显性遗传病)或1/4(隐性遗传病)的发病风险。

(二)群体发病率存在种族(民族)差异

不同的种族(民族)有不同的遗传基础,发病率的种族(民族)差异表明多基因遗传病存在遗传基础。比如,在日本,先天性畸形足的发病率是1.4%,而美国则高达5.5%。

(三)发病率与患者亲属级别(亲缘系数)有关

与患者亲缘系数相同的亲属有相同的发病风险。例如,患者的同胞、子女均为一级亲属,亲缘系数为0.5,这些个体的发病风险相同。

随着亲属级别的降低,患者亲属的发病风险迅速降低,群体发病率越低的疾病,这种趋势越明显(表6-3)。例如,与群体发病率相比,唇裂+腭裂患者的一卵双生同胞发病风险增高400倍,而患者一级亲属的发病风险则迅速下降,仅为群体发病率的40倍。这种随着亲属级别的降低,患者亲属发病率迅速下降的趋势从易患性变异的正态分布曲线

上看得更直观(图 6－5)。

表 6－3　某些多基因遗传病(畸形)患者不同级别亲属的发病风险比较

疾　病	群体发病率	患者亲属的发病风险(群体发病率的倍数)			
		一卵双生	一级亲属	二级亲属	三级亲属
唇裂、唇裂＋腭裂	0.001	×400	×40	×7	×3
足内翻	0.001	×300	×25	×5	×2
神经管缺损	0.002		×8		×2
先天性髋关节脱臼	0.002	×200	×25	×3	×2
先天性幽门狭窄	0.005	×80	×10	×5	×1.5

(四)近亲婚配对发病率的影响不大

近亲婚配时，子女从亲代获得相同的致病基因可能性增高，从而增高发病风险，但其效应不像常染色体隐性遗传病那样显著。

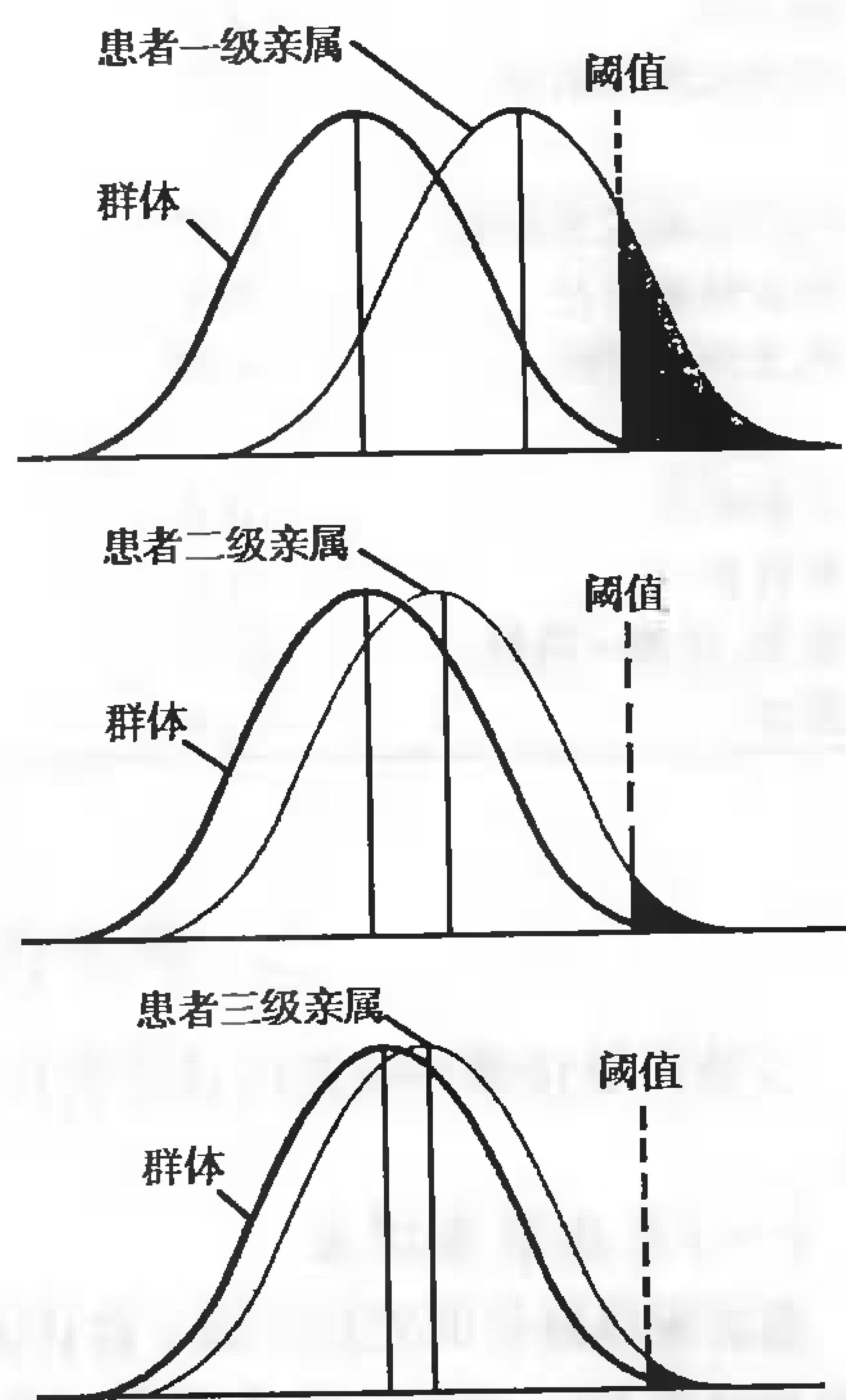


图 6－5　群体和患者不同级别亲属发病率比较

四、多基因遗传病再发风险的估计

当一个家庭中出现某种多基因遗传病患者时，家庭中其他成员患这种多基因遗传病可能性就是多基因遗传病的再发风险。由于多基因遗传病涉及多种遗传因素和环境因素，发病机制复杂，故很难像分析单基因遗传病那样，准确推算出发病风险。一般来说，多基因遗传病患者的一级亲属的发病风险与下列因素有关。

(一)遗传率与群体发病率

如果 2 种多基因遗传病的群体发病率相同，遗传率高的多基因遗传病，患者一级亲属发病率也高；如果 2 种多基因遗传病的遗传率相同，群体发病率高的多基因遗传病，患者一级亲属发病率也高。一级亲属发病率和遗传率、群体发病率的关系可用图 6－6 表示。图中横坐标为群体发病率，斜线为遗传率，纵坐标为患者一级亲属发病率。例如，已知原发性高血压的群体发病率为 6%，遗传率为 62%，查图可知，患者一级亲属发病率为 16%。如果知道了某种多基因遗传病的群体发病率和患者一级亲属发病率，可用图解法求得遗传率的粗略估计值。例如，已知先天性房间隔缺损的群体发病率为 0.1%，患者一级亲属发病率为 3.3%，查图可知先天性房间隔缺损的遗传率约为 74%。

相当多的多基因遗传病的群体发病率为 0.1%～1%，遗传率为 70%～80%，在这种情况下，患者一级亲属发病率(f)近似于群体发病率的平方根($\sqrt{p}$)，即 $f=\sqrt{p}$，这就是 Edward 公式。在此情况下，应用公式比图解法更方便。例如，我国人群中，唇裂＋腭裂的群

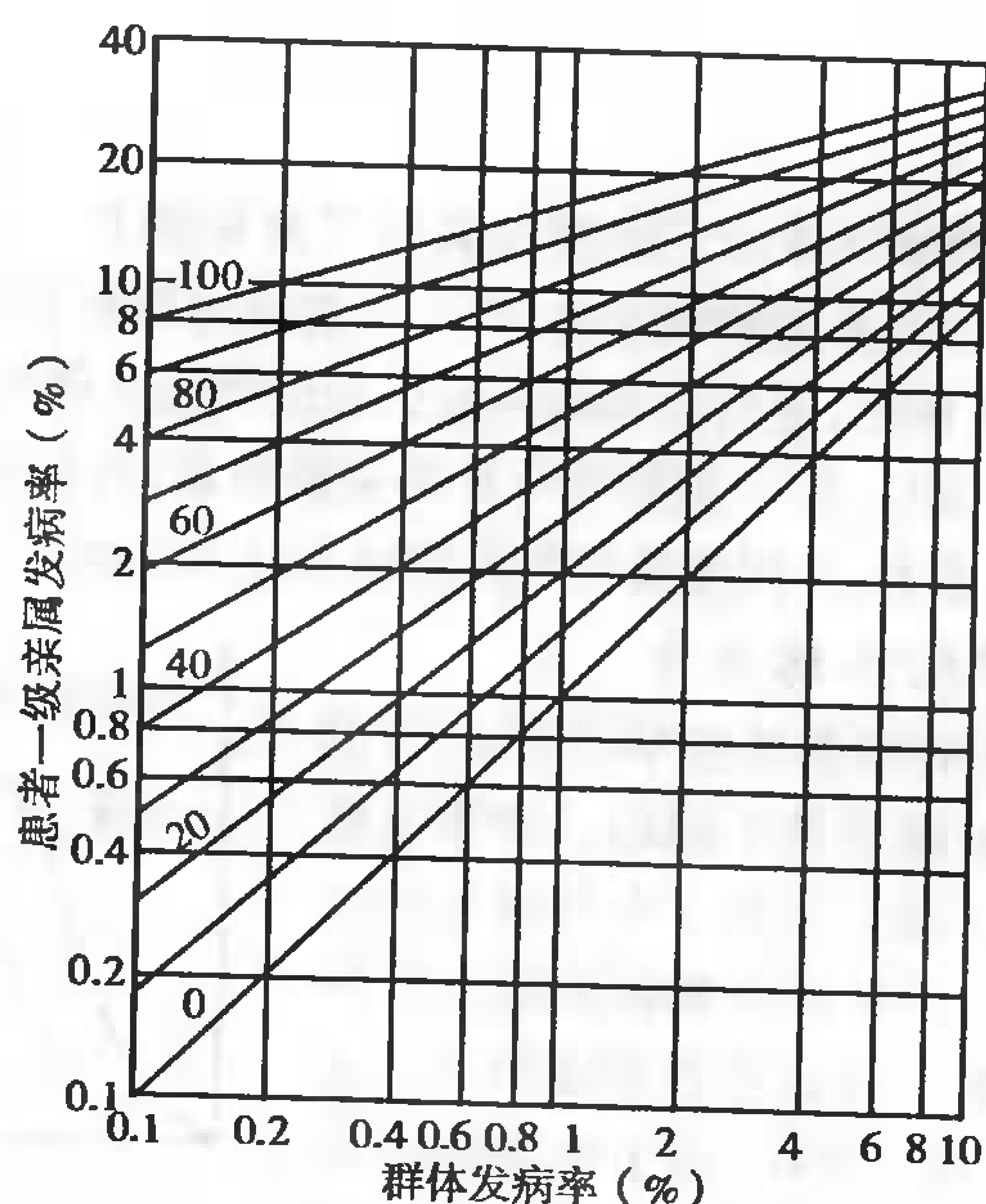


图6-6 患者一级亲属发病率与遗传率、群体发病率的关系
[斜线代表遗传率(%)]

体发病率为0.17%，遗传率为76%，运用公式计算：$f=\sqrt{p}=\sqrt{0.17\%}\approx4\%$，即患者一级亲属发病率约为4%。值得注意的是，应用Edward公式必须具备2个条件，即多基因遗传病的群体发病率为0.1%～1%，遗传率为70%～80%，否则，只能用图解法。再以上文提到原发性高血压为例，该多基因遗传病不具备应用Edward公式的条件，如按Edward公式计算，$f=\sqrt{p}=\sqrt{6\%}\approx24.5\%$，这与根据图解法得到的正确结果($f=16\%$)相差甚远。

（二）家庭中已患病人数

一般来说，一个家庭中已患病人数越多，该病的再发风险越大，这是由于基因的累加效应所致。例如，一对夫妇已有1个唇裂患儿，再次生育唇裂患儿的风险为4%；若果真又生出1个这样的患儿，则表明夫妇二人都带有比原先估计更多的致病基因，虽然他们均未发病，但其易患性极其接近阈值。此时，他们生育第3个唇裂患儿的风险将增加2～3倍，即近于10%。Smith曾研制了一个表格(表6-4)，根据双亲和同胞中已患病的人数估

表6-4 根据家庭中患病人数估计多基因遗传病再发风险(%)

群体发病率(%)	遗传率(%)	双亲患病数 0			1			2		
		患病同胞数			患病同胞数			患病同胞数		
		0	1	2	0	1	2	0	1	2
1.0	100	1	7	14	11	24	34	63	65	67
	80	1	6	14	8	18	28	41	47	52
	50	1	4	8	4	9	15	15	21	26
0.1	100	0.1	4	11	5	16	26	62	63	64
	80	0.1	3	10	4	14	23	60	61	62
	50	0.1	1	3	1	3	9	7	11	15

计再发风险。

(三)病情严重程度

多基因遗传病的基因累加效应还表现在病情严重程度上。与病情较轻的患者相比,病情严重的患者必定带有更多的致病基因,其父母也必定带有更多的致病基因,易患性更接近阈值。所以,再次生育时,生出患儿的风险也相应增高。例如,只有一侧唇裂的患者,其同胞的发病风险为2.46%;而一侧唇裂合并腭裂的患者,其同胞的再发风险为4.21%;而双侧唇裂合并腭裂的患者,其同胞的再发风险则高达5.74%。

(四)群体发病率的性别差异

当某种多基因遗传病的群体发病率存在性别差异时,说明不同性别的阈值是不同的。群体发病率低的性别阈值高,故该性别患者子女发病风险也高,这称为Carter效应。这是因为群体发病率低的性别患者,必然携带比另一性别患者更多的致病基因,才能达到阈值而发病。因此,其子女获得较多致病基因的风险就会增大,尤其是与其性别相反的后代。例如,先天性幽门狭窄的群体发病率,男性为0.5%,女性为0.1%,即男性的群体发病率是女性的5倍,也就是说女性的阈值高于男性。男性患者儿子发病率是5.5%,女儿发病率是2.4%;女性患者儿子发病率高达19.4%,女儿发病率也达到了7.3%。在这里,Carter效应是非常明显的(图6-7)。与此相反,女性强直性脊椎炎的群体发病率是男性的5倍,而男患者子女的发病率远高于女性患者的子女(表6-2)。

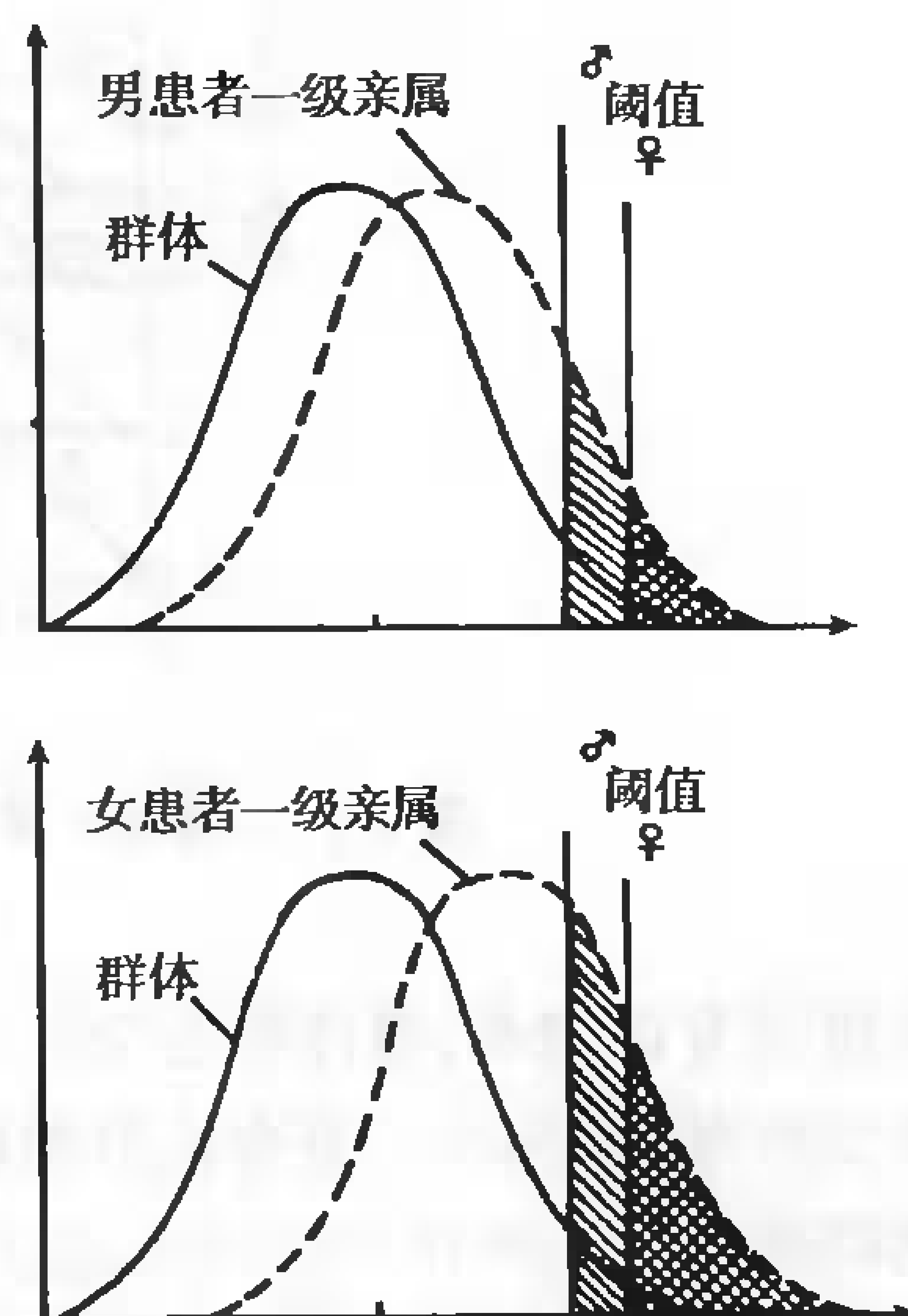


图6-7 群体与患者一级亲属发病率的性别差异比较

第三节 多基因遗传病的研究进展

前已述及,多基因遗传病属于复杂疾病,就遗传因素而言,它们受控于多对微效基因背景上的易感主基因(major susceptibility gene),简称易感基因,发现这些易感基因是认识复杂疾病病因的关键。目前,多基因遗传病易感主基因的研究主要从两方面进行探索:一方面是收集家系资料,用统计学方法进行分类分析、优势对数计分法连锁分析、患病同胞对分析、群体关联分析等来证实主基因的存在;另一方面,用候选基因检测法或用遗传标记来定位易感主基因,并用定位克隆法来鉴定这些易感基因。以下主要介绍近年来常见多基因遗传病的研究进展。

一、原发性高血压

原发性高血压遗传度为40%~60%,识别和克隆原发性高血压主易感基因将从根本上阐明原发性高血压的遗传本质和发病机制,进而对该病的临床个体化治疗、预后判断、

易者的早期检出及预防产生重大影响。目前已经研究过的原发性高血压候选基因超过70个，其中受到广泛关注的有以下几个。

(一)血管紧张素原(AGT)基因

Eunemaitre等用受累同胞对连锁分析法，发现*AGT*基因座位与高血压之间存在紧密连锁关系，以后各国不同种族的研究证实了这一发现。他们还发现，*AGT*基因存在15种变异体，其中M235T变异体(即基因突变导致第235号氨基酸由甲硫氨酸转变为苏氨酸)与高血压相关联，235T纯合个体的血浆AGT水平显著高于235M纯合个体。最近一项研究还表明，TT、TM基因型携带者的收缩压、舒张压和血浆AGT浓度均显著高于MM基因型携带者。但是不少报道无法证实*AGT*基因M235T多态性与高血压有关，因此，导致高血压的AGT基因缺陷还有待进一步研究，应注意在*AGT*基因的调控序列中筛查基因突变。此外，原发性高血压相关基因在*AGT*基因座位附近的可能性也不能排除。

(二)血管紧张素转化酶(ACE)基因

人*ACE*基因在第16内含子中存在插入型(insertion,I)、缺失型(deletion,D)多态性。Soubrier发现血清ACE水平与*ACE*基因多态性之间有显著相关性，血清ACE活性高低在不同*ACE*基因型人群中依次为DD > DI > II型。目前一般认为，*ACE*基因多态性与高血压关系不大，但*ACE*基因多态性与心、脑、肾等靶器官损害有关。

(三)α - *adducin* 基因

Bianchi等在米兰高血压大鼠(MHS)中发现一种细胞骨架蛋白α - adducin的基因存在错义突变，此突变可激活肾小管细胞膜上的$Na^+ - K^+$泵，从而影响肾小管对钠的重吸收。进一步研究发现，人的α - *adducin*基因突变与高血压间也存在一定的相关关系，而且，α - *adducin*突变基因携带者对利尿剂的降压作用优于对照组。α - *adducin*基因突变是否为盐敏感性高血压的遗传基因，是否还存在与盐敏感有关的其他基因，值得进一步研究探讨。

(四)B型利钠肽(NP)受体基因

NP在原发性高血压的发生中起决定性作用。C型NP通过B型*NP*受体(NPR - B)起扩张动脉、降低血压、抑制血管平滑肌细胞增殖的作用。*NPR - B*基因长约16.5 kb，由22外显子构成，外显子2 - 内含子接头下游约150 bp处存在一个CA/GT(11)微卫星重复序列，原发性高血压患者的GT(11)的频率为0.316(65/206)，正常血压组为0.218(44/202)，二者之间的频度差别明显，多变量逻辑回归分析提示*NPR - B*基因型与原发性高血压有明显关联(OR 1.5;95% 1.02 ~ 2.35)，提示*NPR - B*基因或与之关系密切的基因或区域的突变是原发性高血压的一个病因。

二、冠心病

华中科技大学人类基因组研究中心的王擎等人对一个冠心病的大家系进行了研究，该家系共有21成员，拥有13名冠心病患者，其中9人曾发生过心肌梗死。研究结果证实，*MEF2A*基因的一种突变导致了冠心病和心肌梗死的发生。该家系患者体内*MEF2A*基因都存在突变，而家族中健康成员携带的是该基因的野生型。*MEF2A*基因的发现为研

究冠心病和心肌梗死发病机制，开发预防和治疗该疾病的新药具有重要价值。新发现的 *MEF2A* 基因编码的调控蛋白，对血管内皮细胞中成百上千基因的表达起到控制作用。该基因编码的蛋白在正常情况下有 499 个氨基酸。王擎等人在研究中发现，*MEF2A* 基因的一种缺失突变会使调控蛋白损失 7 个氨基酸，这一突变有可能破坏其他基因表达程序，使冠状动脉血管发育形成缺陷，导致巨噬细胞等更容易侵人血管，产生动脉粥样硬化斑块堵塞血管，阻碍血液流入心脏，最终造成心脏组织缺氧甚至坏死，患者可能因此而发生心肌梗死。

三、糖尿病

糖尿病有明显的遗传倾向并存在显著遗传异质性，除少数是由于单基因突变所致外，大部分是由多基因及环境因子共同参与及相互作用引起的多基因遗传病。

胰岛素依赖性糖尿病(insulin dependent diabetes mellitus，IDDM)又称为Ⅰ型糖尿病，遗传易感基因十分复杂，并与 *HLA* 基因复合体密切相关。研究表明，单元型 HLA - A1、C1、B56、DR4、DQ8 有非常高的绝对危险性，而近 50% 的遗传危险性可归于 *HLA* - Ⅱ类基因(DR、DQ、DP)，当 HLA - DQ B1 链 57 位为非天门冬氨酸(为天门冬氨酸时是保护基因)、HLA - DQ A1 链为 52 位精氨酸时易感Ⅰ型糖尿病。近年来，筛选出的第 2 代 IDDM 易感基因有：*IDDM* 2(11p15)，*IDDM* 3(15q26)，*IDDM*4(11q13)，*IDDM*5(6q25)，*IDDM*7(2q31)，*IDDM*8(6q27)，*IDDM*11(14q24.3 - q31)，*IDDM*12(2q33 上的 CTLA4)，*IDDM*13(2q34)，*GCK*3(葡萄糖激酶 3,7p)。

非胰岛素依赖性糖尿病(non - insulin dependent diabetes mellitus，NIDDM)又称为Ⅱ型糖尿病，发病机制更为复杂，相关基因有：胰岛素基因、胰岛素受体基因、胰岛 β 细胞葡萄糖转运蛋白 2 基因、糖原合成酶基因、胰高血糖素受体基因、磺脲受体基因等。

我国糖尿病研究专家王姮教授经过多年的科研探索，借助微卫星技术，通过分布于人体基因组上的 358 对微卫星引物，对北方地区 32 个汉族Ⅱ型糖尿病家系的 DNA 进行全基因组扫描，再对扫描获得的信息和图谱进行准确的基因分型，初步确定了Ⅱ型糖尿病易感基因位点。上海第二医科大学附属瑞金医院罗敏教授带领的课题组，采用全基因组扫描筛查技术，对上海、山东、福建、辽宁等地至少 2 代人以上都患有Ⅱ型糖尿病的 102 个家系进行分析，已在国际上率先将中国人Ⅱ型糖尿病易感基因定位在 9 号染色体的 9p21 区域，为最后克隆Ⅱ型糖尿病易感基因奠定了基础。

小　结

人类多基因遗传病的发生受多对等位基因控制，环境因素的作用较明显。多对等位基因是共显性的，单个基因的作用微小，但有累加效应，故称为微效基因、加性基因。多基因遗传的性状在群体中的变异分布是连续的，仅有数量差异，称为数量性状。数量性状变异呈正态分布，即以中间类型(均数)为对称轴，两侧对称分布。越远离均数，变异的分布越少，曲线越低。

个体在遗传基础和环境因素共同作用下，患某种多基因遗传病的风险称为易患性，群

体中每个个体的易患性变异呈正态分布。能够使个体发病的易患性限度称为阈值。阈值将群体分为两部分：处于阈值以下的未发病者，达到阈值及阈值以上的患者。在环境因素相同的情况下，阈值代表发病所需的、最低致病基因数。多基因遗传病的发病率就是阈值右侧面积占正态曲线下总面积的比率。

某种多基因遗传病，个体的易患性可根据家族成员的发病情况粗略估计；群体易患性平均值可从群体发病率估计。一种多基因遗传病群体发病率越低，说明其阈值越高，阈值与易患性平均值间距离越大，相对来说，易患性平均值越低。

遗传率，也称遗传度，是指遗传因素在由遗传因素和环境因素共同作用所致多基因遗传病中所起作用的大小。

与单基因遗传病相比，多基因遗传病有不同的遗传特点：一是呈现家族聚集现象；二是群体发病率存在种族（民族）差异；三是患者亲属的发病风险与亲属级别（亲缘系数）有关；四是近亲婚配致子女发病风险增高的效应不像单基因遗传病显著。

与多基因遗传病再发风险有关因素包括：多基因遗传病的遗传率和群体发病率，家庭中患病人数，患病成员的病情严重程度，群体发病率的性别差异，即 Carter 效应。

（蔡绍京）

第七章　染色体病

染色体数目异常或结构畸变统称为染色体畸变(chromosomal aberration),由此产生的疾病称为染色体病(chromosomal disorder)。染色体是遗传物质的载体,人类每条染色体上平均带有1 000多个基因,即使染色体上微小的结构改变,也会导致许多基因的增加或丢失,从而造成许多基因表达和代谢的紊乱。染色体数目异常或结构畸变引起的染色体病,常产生一系列临床表现,故称其为染色体畸变综合征(chromosomal aberration syndrome)。迄今为止,已报道的染色体数目和结构异常类型有10 000余种,已确定的染色体病约100余种,染色体病分为常染色体病和性染色体病2类。

第一节　染色体畸变

一、染色体畸变的诱因

在正常情况下,人体细胞中可有少数染色体畸变发生,一般在1%以下。许多正常的细胞系经过数代培养后才可看到有些细胞染色体数目、结构发生变化,有人认为这是因为在培养过程中细胞核通过形成微核而丢失部分染色体,或是细胞间发生融合后又经微核的丢失而成为染色体数目增多的非整倍体。但在某些化学物质、射线和病毒等因素作用下,人类细胞中染色体异常频率大大增高。影响染色体异常的因素大多为环境因素,但也可能是遗传因素所致。

(一)物理因素

各种射线,如X射线、γ射线、α和β粒子、中子等可诱发染色体畸变。射线能破坏DNA链之间的共价结构,导致染色体断裂。长期接受射线治疗或从事放射工作的人员,由于长期小剂量的累积作用,可引起体细胞或生殖细胞染色体畸变。染色体畸变率随射线剂量的增高而增高。

(二)化学因素

许多药物可导致染色体畸变。妊娠早期使用药物不当可引起胎儿染色体畸变,某些抗癌药物如环磷酰胺、甲氨蝶呤等也可引起染色体畸变,有机磷类农药、工业毒物、食品添加剂、防腐剂、保鲜剂等均可诱发染色体畸变。

(三)生物因素

霉菌产生的毒素,例如黄曲霉素、柄曲霉素和棒曲霉素等,有诱发染色体畸变的作用;各种病毒,如SV40病毒、仙台病毒、牛痘病毒、风疹病毒、带状疱疹病毒、麻疹病毒、肝炎病毒、流行性腮腺炎病毒等可诱发宿主细胞染色体畸变。Rous肉瘤病毒、Buikitt淋巴肉瘤病毒等不仅可引起染色体畸变,还有致癌作用。

(四)机体内在因素

母亲的生育年龄与21三体综合征患儿的出生率有关,其他三体型综合征也有类似的情况。一般认为,母亲生育年龄越晚,生殖细胞在体内停留的时间越长,受到不利因素影响的机会越多,在减数分裂时,越容易产生染色体不分离而导致染色体数目畸变。染色体畸变的发生与胚胎发育的宫内环境也有关系。遗传素质也与染色体畸变有关,一些常染色体隐性遗传病患者常发生染色体自发断裂。

二、染色体数目异常

正常人体细胞具有46条染色体,为二倍体(diploid)。二倍体细胞中每种常染色体均为2条,同时有XX或XY性染色体。任何偏离这一限定染色体数目的改变都称为染色体数目异常(chromosomal numerical abnormality)。少数情况下,体细胞内有2种以上染色体数目异常,但增减的染色体数目相抵,这种染色体数目异常称为假二倍体(pseudodiploid)。例如,某女性的核型为46,XX,-18,+21,这表示染色体总数是46,但少1条18号染色体,多1条21号染色体。

(一)整倍体

具有完整染色体组数的细胞称为整倍体(euploid)。与正常二倍体细胞相比,整倍体异常的细胞内染色体组增加,形成多倍体(polyploid)。按照细胞内染色体的组数,多倍体分为三倍体(triploid)、四倍体(tetraploid)等。在人类,三倍体是致死的,极为罕见,但三倍体在流产胎儿中较常见,是流产的重要原因之一。国外有少数三倍体婴儿的报道,这种婴儿往往在出生后不久即夭亡。三倍体产生的原因一般认为有2种:一是双雌受精,卵子发生的减数分裂过程,由于某种原因未能排出极体,极体的染色体保留在卵母细胞内,结果形成二倍体卵子,二倍体卵子($2n$)与正常精子(n)受精,形成69,XXX或69,XXY三倍体合子;二是双雄受精,即同时有2个精子进入卵细胞参与受精,可形成69,XXX、69,XXY或69,XYY 3种三倍体合子。

全身性四倍体更为罕见,但四倍体和其他多倍体细胞在肝、子宫内膜、骨髓细胞、肿瘤组织和培养细胞中并不罕见。四倍体产生的原因,一是核内复制(endoreduplication)(图7-1),即在1个细胞周期中,染色体复制2次;复制形成的染色体两两平行排列在一起称为双倍染色体(diplochromosome);二是核内有丝分裂(endomitosis),即在细胞分裂中期核膜未能破裂,纺锤体不能形成,无法进行后期的染色单体分离和胞质分裂,造成核内染色体数目加倍。



图7-1　核内复制

(二)非整倍体

细胞中个别染色体的增减形成非整倍体(aneuploid)。少于二倍体者,称为亚二倍体

(hypodiploid)；多于二倍体者，称为超二倍体(hyperdiploid)。非整倍体的形成与减数分裂过程中染色体分离机制异常有关。如果1对同源染色体或2条姐妹染色单体在分裂后期未能分开而同时进入1个子细胞中，这样形成的2个子细胞，1个是超二倍体，1个是亚二倍体。这种情况称为染色体不分离(non-disjunction)。

减数分裂后期Ⅰ染色体不分离，将产生 $n+1$ 型和 $n-1$ 型2种类型配子，且各占一半。异常配子与正常配子结合将分别形成三体型(trisomy)和单体型(monosomy)(图7-2A)。三体型是指细胞中某号染色体增加1条，染色体总数为47($2n+1$)；单体型是指细胞中某对染色体少1条，染色体总数为45($2n-1$)。

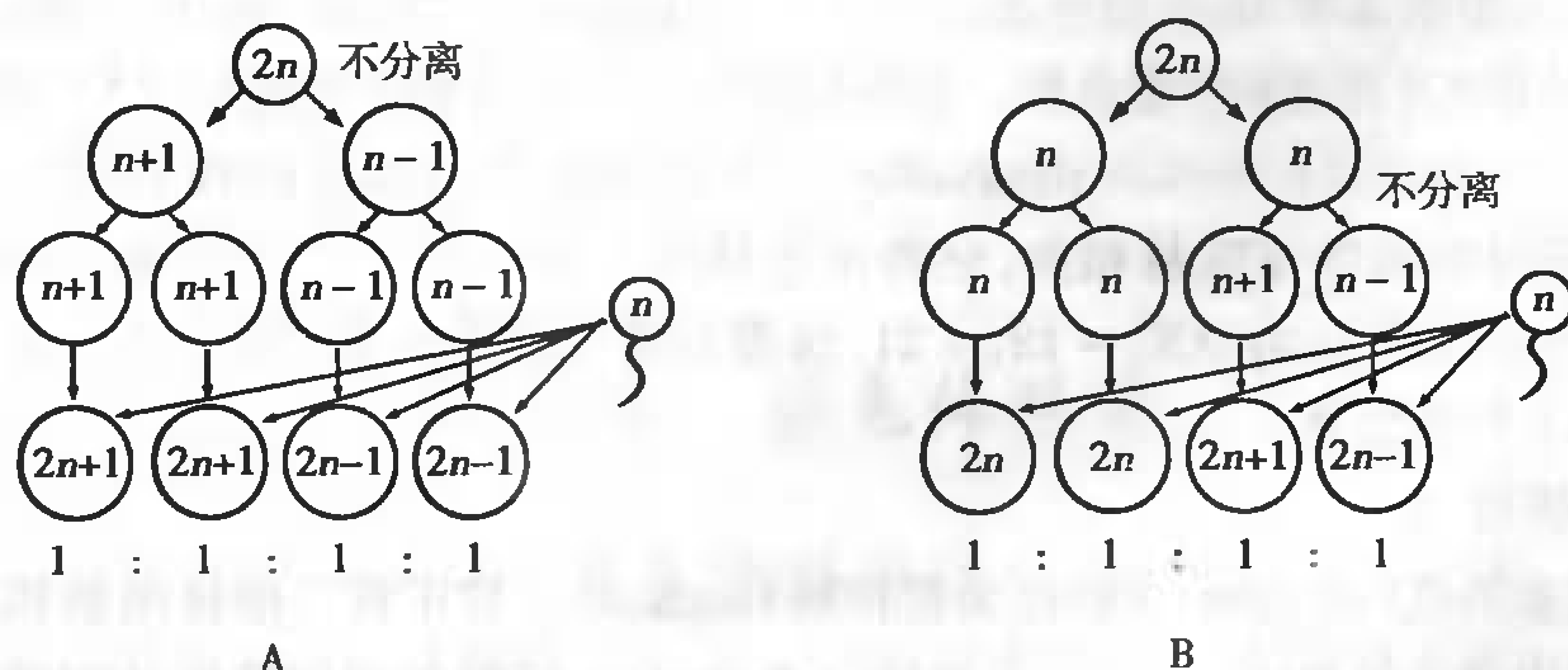


图7-2　减数分裂染色体不分离与非整倍体形成

减数分裂后期Ⅱ染色体不分离，将形成3种不同类型的配子，正常配子占一半，$n+1$ 型和 $n-1$ 型各占1/4。3种类型的配子与正常配子结合后，一半为正常二倍体，三体型和单体型各占1/4(图7-2B)。

三体型和单体型相比，三体型，特别是性染色体的三体型相对多见；单体型少见，其中X染色体单体型相对多见。因为常染色体的单体型，即使是最小的21号、22号染色体的单体型，所造成的整条染色体的基因缺失，将导致基因剂量的严重失衡，干扰细胞的代谢和发育，故仅在流产儿和死婴中见到。在三体型中，较大染色体的三体型个体难以存活，其胚胎发育以流产告终(图7-3)。临床染色体病中最常见的是21三体型和性染色体三体型，其次是18三体型和13三体型。

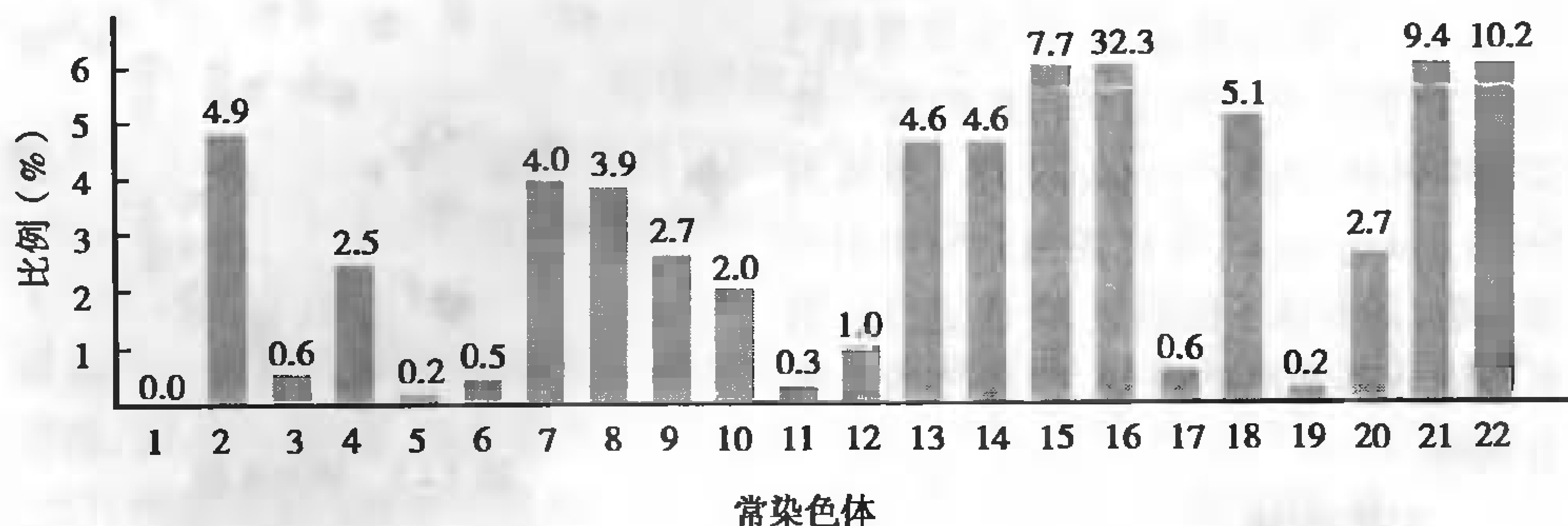


图7-3　常染色体三体自发流产的比率

上述三体型和单体型的产生是减数分裂染色体不分离造成的，而非父母核型异常所致，故称为初级不分离(primary non - disjunction)。如果双亲之一为三体型患者，如母亲为47,XX,+21，其初级卵母细胞中有3条21号染色体，减数分裂过程中，势必形成一种配子细胞中1条21号染色体，另一种配子中2条21号染色体(不分离)，这种由于亲代染色体数目异常导致的减数分裂过程中染色体不分离称为次级不分离(secondary non - disjunction)。上述次级不分离形成的配子与正常精子受精则形成21三体型(47,XX,+21或47,XY,+21)。理论上，21三体型父亲或母亲均可生育21三体型后代，但临床上尚无男性21三体型患者生育子女的报道。

(三)嵌合体

嵌合体是指具有2种或2种以上染色体组成的细胞系的个体。如46,XY/47,XXY、45,X/46,XX/47,XXX等。整倍体和非整倍体可能是指细胞，也可能是指个体，而嵌合体只能是指个体，不存在嵌合体的细胞。细胞系来源于不同合子的嵌合体称为异源嵌合体(chimera)，细胞系来源于同一合子的嵌合体称为同源嵌合体(mosaic)。如果不加特别说明，通常所说的嵌合体均指同源嵌合体。构成嵌合体的细胞可能包含整倍体和非整倍体，也可能均为不同类型的非整倍体，还可能是含有不同结构畸变的同一整倍体($2n$)。嵌合体的形成机制包括有丝分裂不分离和染色体丢失。

有丝分裂不分离(mitotic non - disjunction)发生在受精卵形成后的卵裂过程中，即发生于体细胞中。染色体不分离的方式与减数分裂Ⅱ不分离相同，是姐妹染色单体的不分离，由此产生的嵌合体中各种细胞系的类型及比例取决于不分离发生的早晚。如果发生于受精卵($2n=46$)形成后的第1次卵裂中，将形成三体型(47)和单体型(45)2个细胞系构成的嵌合体，两系各占50%；如果发生在第2次卵裂中，将形成正常二倍体(46)、三体型(47)和单体型(45)3个细胞系(46/47/45)构成的嵌合体，前者占50%，后二者各占25%(图7-4)。这种不分离发生的越晚，正常二倍体细胞系所占比例就越大，异常细胞系的比例就越少，临床表现也就相对越轻。如果染色体不分离发生在第3次卵裂时，二倍体细胞将占75%，而三体型(47)和单体型(45)只各占12.5%。

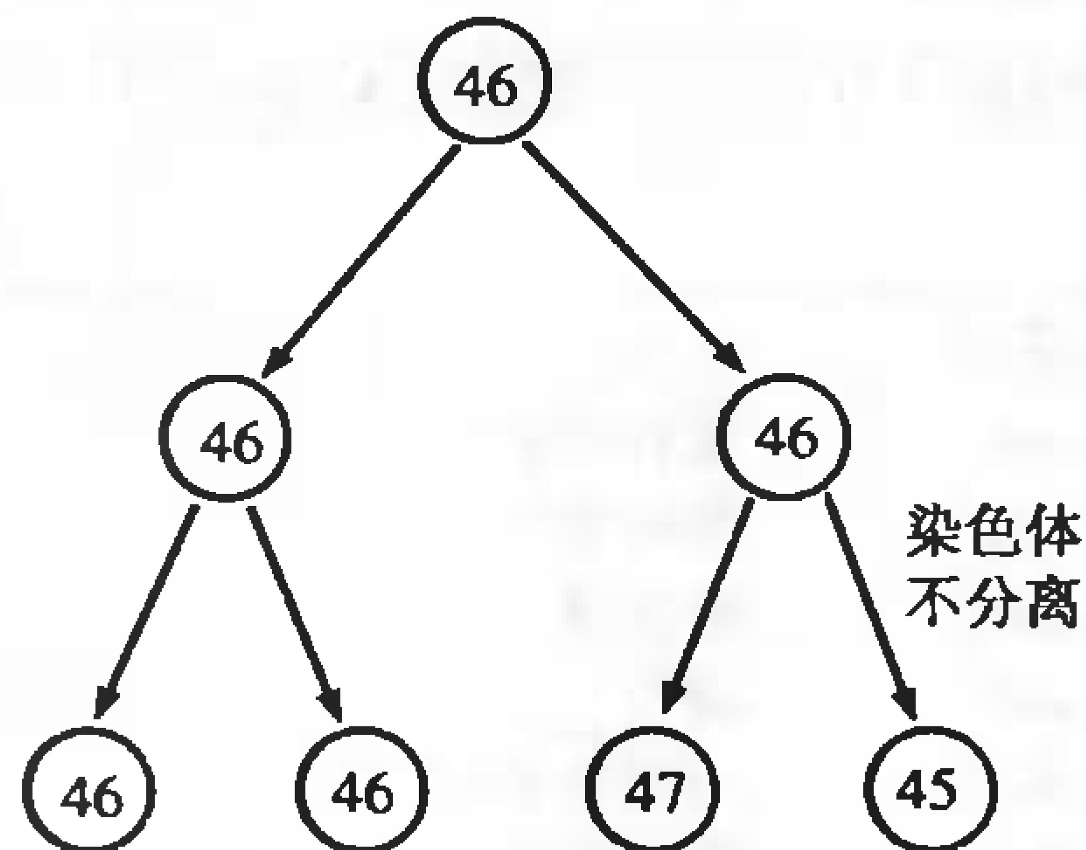


图7-4 有丝分裂不分离与嵌合体形成

染色体丢失(chromosome loss)又称后期延迟(anaphase lag)，发生在细胞分裂的中期至后期阶段，由于某条染色体的着丝粒未能与纺锤丝相连，或染色体在后期移动迟缓，导致某条染色体未进入新的细胞核而滞留在胞质中。滞留在胞质中的染色体最终被分解丢失，导致分裂后一个子细胞缺少1条染色体成为单体型。例如，细胞分裂后期性染色体的丢失可形成45,X/46,XX嵌合体(图7-5)。

三、染色体结构畸变

染色体结构畸变(chromosomal structural aberration)是指染色体部分片段的缺失、重复

或重排。染色体断裂(breakage)和(或)断裂后的异常接合是染色体结构畸变的基础。染色体片段的接合(reunion)只发生在染色体的断端,断端能被再接的特性称为“黏性”。一次断裂产生的2个黏性末端重新接合,不产生遗传效应。同一条染色体或不同染色体上至少2处断裂是异常接合的前提,异常接合将导致染色体结构畸变。异常接合的各种结构异常的染色体称为衍生染色体(derivative chromosome)。衍生染色体是染色体原发性重排的产物,在减数分裂过程中不产生新的重排。如果衍生染色体在减数分裂过程中移位的片段与正常位置的相应片段发生交换,则形成重组染色体(recombined chromosome)。重组染色体是染色体继发性重排的产物(参见本章第四节)。

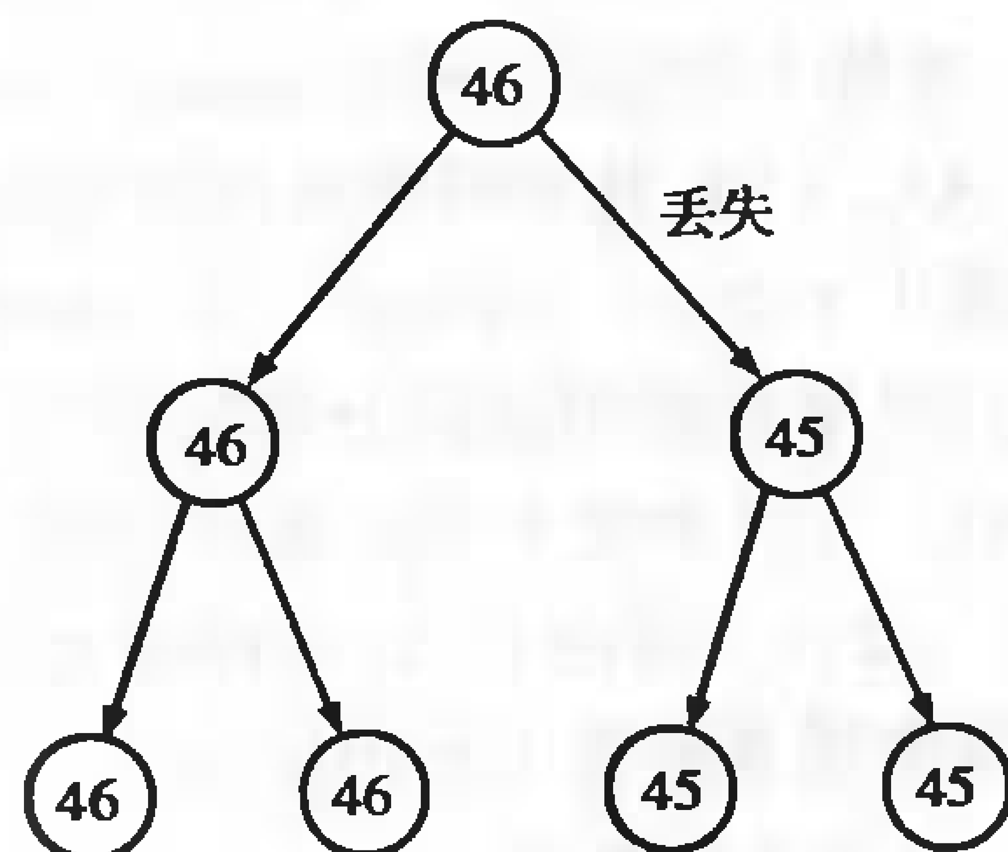


图7-5 有丝分裂染色体丢失与嵌合体形成

(一)染色体畸变的描述方法

为了能够统一规范地描述各种染色体畸变,从1960年以来,历次国际人类细胞遗传学会议都做出了具体规定。人类细胞遗传学命名国际体制(1995)[An International System for Human Cytogenetic Nomenclature(1995),ISCN(1995)],是一个包括自1960年以来8次国际人类细胞遗传学会议上所制定的国际法规。该法规具体规定了核型描述的符号、术语(表7-1)和描述方法等。根据ISCN(1995)的规定,畸变染色体的描述有简式和详式2种方法。

表7-1 核型分析常用符号和术语

符号	含义	符号	含义
ace	无着丝粒片段	rea	重排
cen	着丝粒	rob	罗伯逊易位
chr	染色体	s	随体
del	缺失	sct	次缢痕
der	衍生染色体	t	易位
dic	双着丝粒	tel	端粒
dup	重复	ter	末端
end	核内复制	fem	女性
fra	脆性位点	mal	男性
h	异染色质次缢痕	mat	母源
i	等臂染色体	pat	父源
ider	等臂衍生染色体	?	染色体或染色体结构未能确定
ins	插入	/	隔开嵌合体的不同细胞系
inv	倒位	:	断裂
mar	标记染色体	()	括号内为结构畸变的染色体
mos	嵌合体	::	断裂后重接
p	短臂	,	区分染色体数目、性染色体和染色体异常
Ph	费城染色体	→	从……到
q	长臂	;	分开结构重排的染色体
r	环状染色体	+ 或 -	增加或缺失
rcp	相互易位	+	增加(增长)

1.简式　用简式描述时,染色体的结构改变只用断裂点(break point)表示,简式包括5项:染色体总数、性染色体组成、结构畸变的符号、受累染色体序号和断裂点的臂、区、带号,后2项均分别写在括号内。例如,46,XX,del(5)(p15)表示染色体总数为46的女性患者,5号染色体短臂1区5带断裂,断裂点远端丢失。如果是染色体数目异常,简式描述时,只需描述前2项,但如有需要,在第2项后要描述染色体的增减情况。例如,47,XY,+21表示21三体型男性,46,XX,+18,-21表示18三体型和21单体型假二倍体女性。

2.详式　详式与简式的区别是:第2个括号内不是只描述断裂点的位置,而是描述重排染色体带的构成。如上例46,XX,del(5)(p15),用详式描述则为46,XX,del(5)(qter→p15:)。

(二)染色体结构畸变的主要类型

1.缺失(deletion)　染色体臂的部分丢失称为缺失。分为中间缺失和末端缺失2种类型。染色体臂上发生1次断裂,断端以远片段丢失,称为末端缺失(terminal deletion);染色体臂上发生2次断裂,中间片段丢失,称为中间缺失(intercalary deletion)。缺失造成细胞内该片段实际上只剩下另一同源染色体上一份,故又将缺失称为部分单体型(partial monosomy)。

例如:1号染色体q21断裂形成末端缺失(图7-6),简式为46,XX,del(1)(q21),详式为46,XX,del(1)(pter→q21:)。

3号染色体q21和q25处断裂形成中间缺失(图7-7),简式为46,XX,del(3)(q21q31),详式为46,XX,del(3)(pter→q21::q25→qter)。

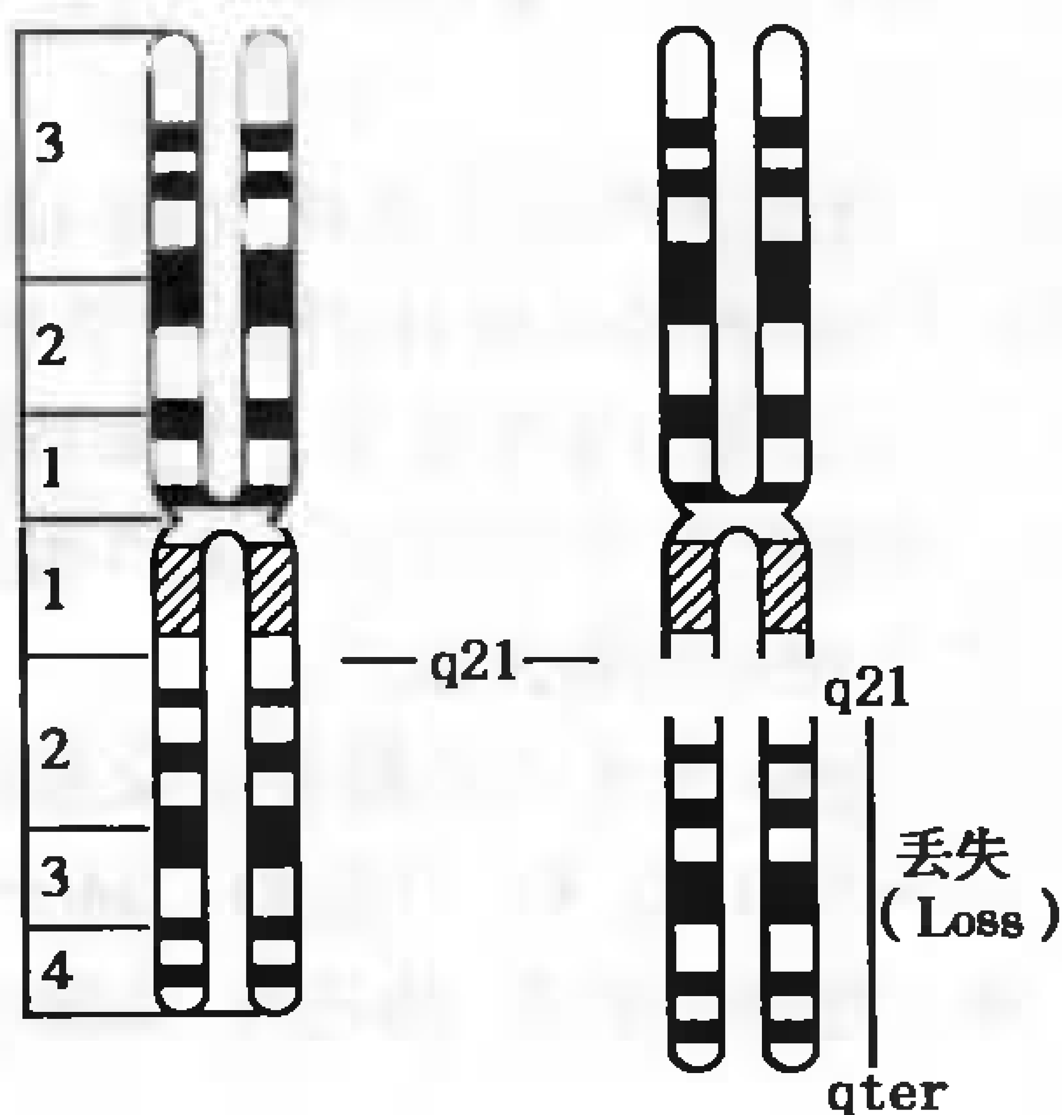


图7-6　1号染色体末端缺失

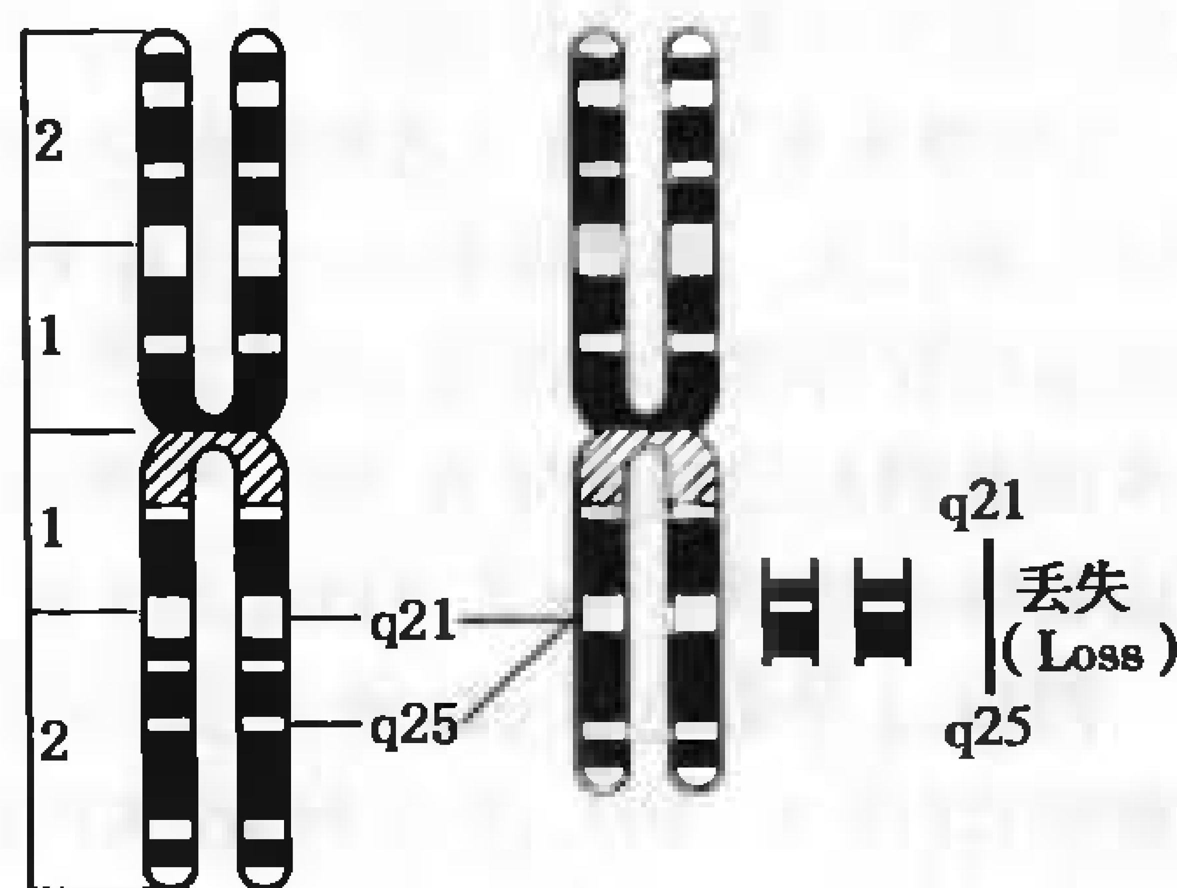


图7-7　3号染色体中间缺失

2.倒位(inversion)　染色体发生2处断裂,中间的片段倒转180°后重新接合,称为倒位。发生在同一臂(长臂或短臂)内的倒位称为臂内倒位(paracentric inversion);如果断裂分别发生在长臂和短臂,倒位的部分包括着丝粒,则称为臂间倒位(pericentric inversion)。

例如:1号染色体p22和p34处断裂形成臂内倒位(图7-8),简式为46,XY,inv(1)(p22p34),详式为46,XY,inv(1)(pter→p34::p22→p34::p22→qter)。

4号染色体p15和q23处断裂形成臂间倒位(图7-9),简式为46,XY,inv(4)(p15q21)详式为46,XY,inv(4)(pter→p15::q23→p15::q23→qter)。

3.重复(duplication)　染色体臂上某一节段有2份或2份以上,称为重复。重复大多

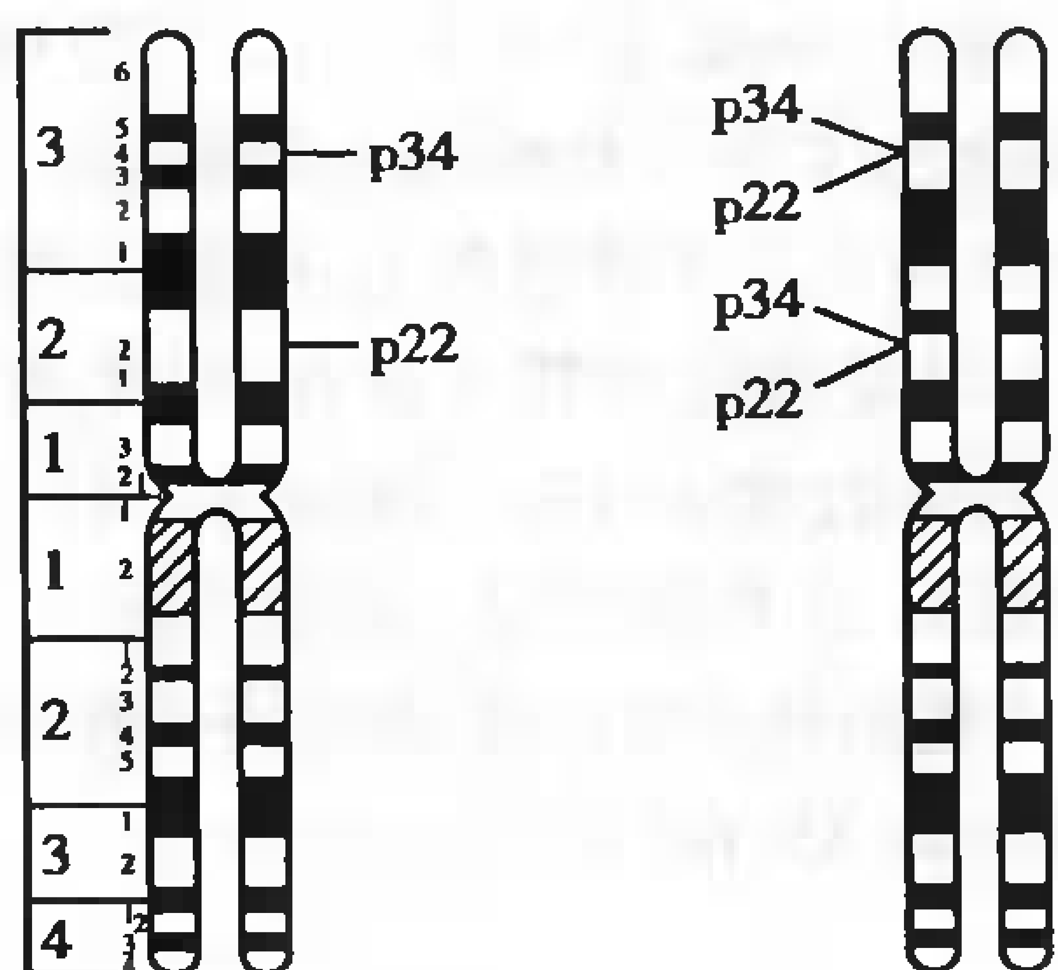


图 7-8　1 号染色体短臂臂内倒位

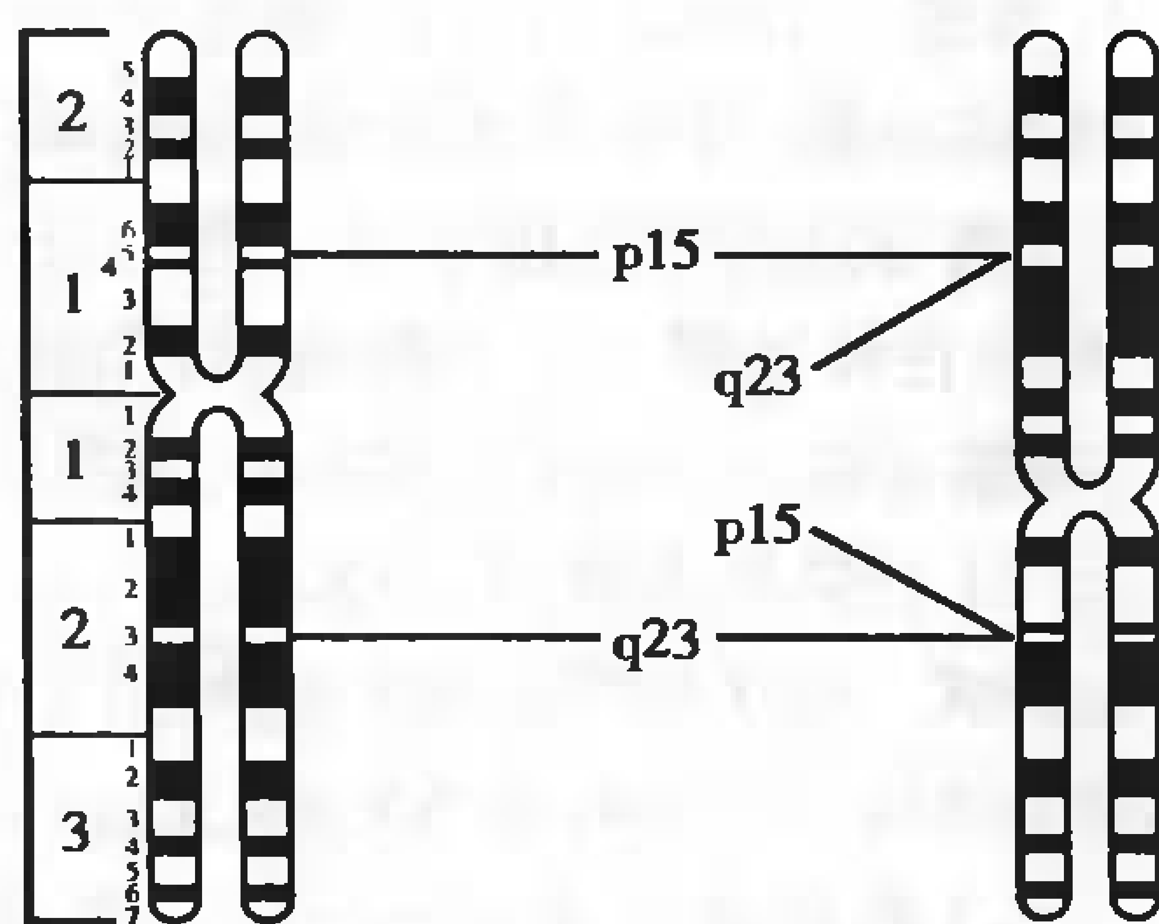


图 7-9　4 号染色体臂间倒位

是减数分裂时同源染色体间错误配对和不等交换所致。染色体上重复 1 份的片段，在该细胞内实际上存在 3 份，故又称为部分三体型(partial trisomy)。

4. 易位(translocation)　染色体片段位置的改变称为易位。发生在 1 条染色体内易位的称为染色体内易位(intrachromosomal translocation)；发生在 2 条同源或非同源染色体间的易位称为染色体间易位(interchromosomal translocation)。染色体间易位又分为非相互易位(nonreciprocal translocation)和相互易位(reciprocal translocation)。

(1)非相互易位　又称转位(transposition)。是指染色体断裂后，一条染色体的片段移位接合到另一条染色体的断端上，另一条染色体的片段丢失。

(2)相互易位　指 2 条染色体同时断裂，两无着丝粒片段互换位置重新接合，形成 2 条衍生染色体。相互易位是比较常见的染色体结构畸变类型，各号染色体间都可发生。相互易位仅有位置的改变，没有染色体上遗传物质的增减，故又称为平衡易位。平衡易位通常没有明显的临床异常，然而平衡易位携带者与正常人婚配生育的子女有可能得到重组染色体，导致某一易位节段增加 1 份(部分三体型)，并产生相应的临床表现。

例如：2 号染色体 q21 和 5 号染色体 q31 处断裂，2 条染色体的无着丝粒片段交换后重新接合(图 7-10)，简式为 46，XY，t(2；5)(q21；q31)，详式为 46，XY，t(2；5)(2pter→2q21：：5q31→5qter；5pter→5q31：：2q21→2qter)。涉及 2 条染色体的畸变，描述时，性染色体及序号靠前的染色体先描述。

(3)罗伯逊易位(Robertsonian translocation)　又称为着丝粒融合(centric fusion)，是发生在 2 条近端着丝粒染色体(D/D，D/G，G/G)之间的一种特殊形式的相互易位。根据 ISCN(1995)规定，着丝粒区定义为 10，着丝粒的短臂侧部分为 p10，着丝粒的长臂侧部分为 q10。罗伯逊易位的断裂点为 q10，断裂后，两长臂部分接合形成一条大的染色体，两短臂部分接合形成的小染色体不稳定，易丢失。

例如：14 号和 21 号染色体间的罗伯逊易位(图 7-11)，简式为 46，XY，der(14；21)(p10；q10)，+21，详式为 46，XY，der(14；21)(14qter→14p10：：21q10→21qter)，+21。这表示患者 14 号和 21 号 4 条染色体中的 2 条形成了 1 条 14/21 衍生染色体，另外，还多了 1 条 21 号染色体。

5. 环状染色体(ring chromosome)　染色体的长臂和短臂远端各有 1 次断裂，有着丝粒的主体部分首尾接合形成环状染色体，无着丝粒片段通常在细胞分裂过程中丢失。因此，环状

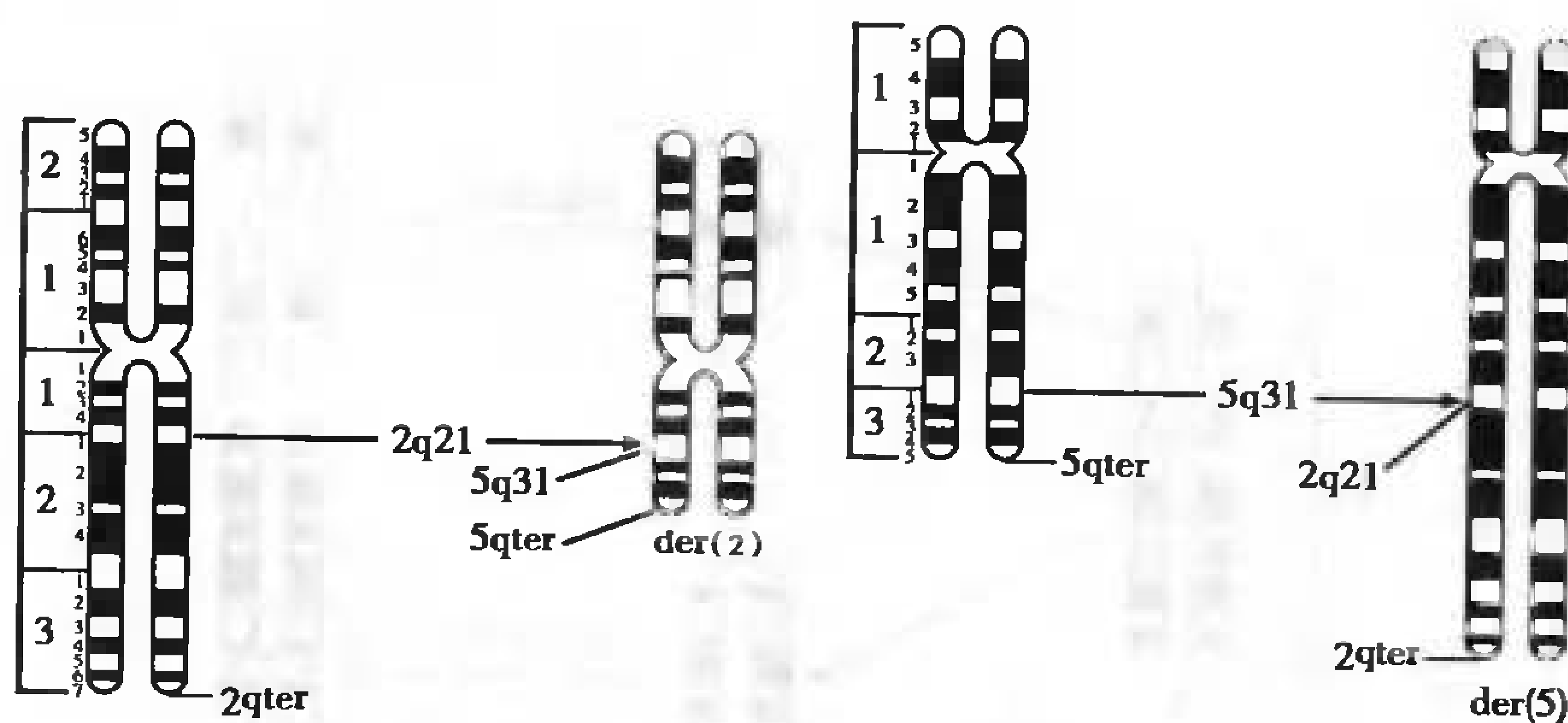


图 7-10 2号与5号染色体间相互易位

染色体实际上同时存在染色体末端缺失。细胞受到辐射损伤,特别是在恶性肿瘤患者放射治疗时,细胞中常见到环状染色体。

例如:2号染色体 p21 和 q31 断裂形成环状染色体(图 7-12),简式为 46,XY,r(2)(p21q31),详式为 46,XY,r(2)(::p21→q31::)。

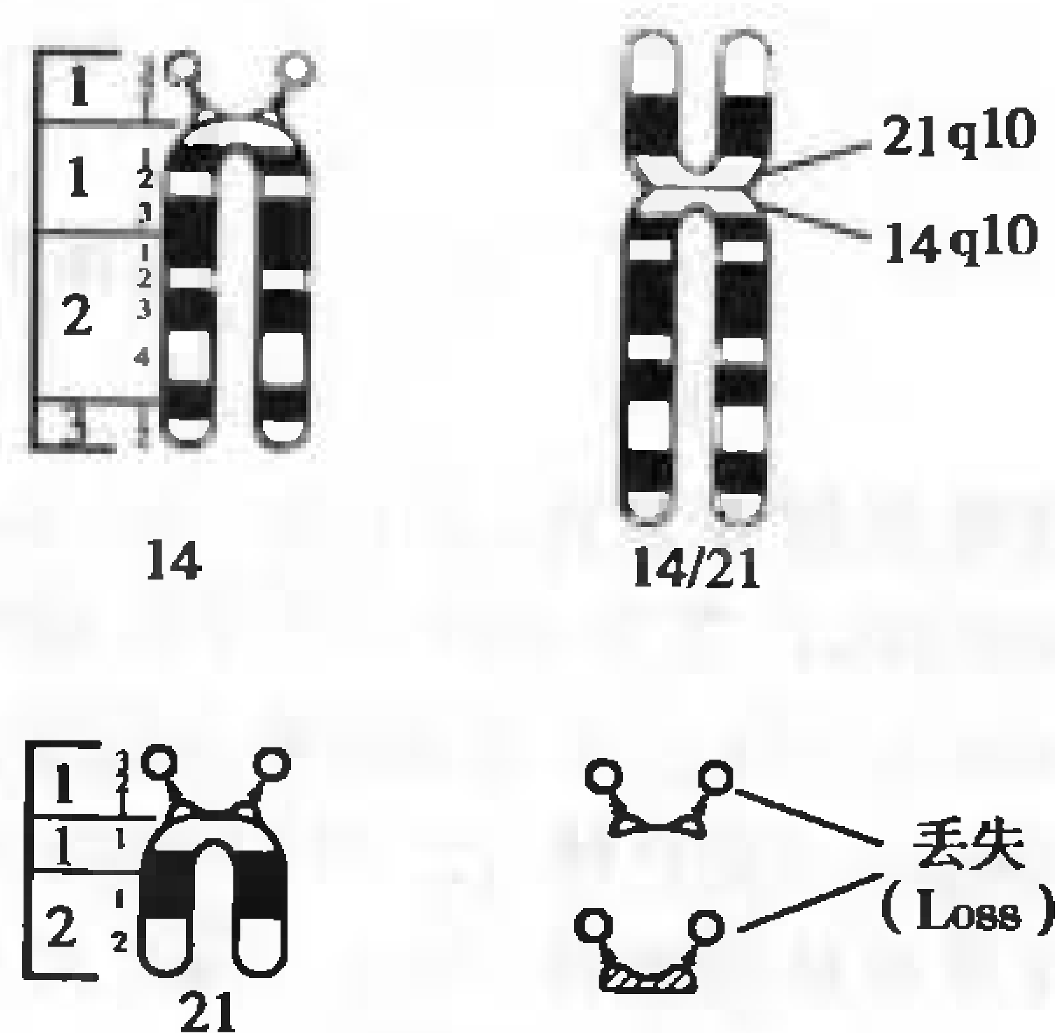


图 7-11 14/21 染色体罗伯逊易位

6. **等臂染色体(isochromosome)** 形态和遗传组成上完全相同的染色体称为等臂染色体。一般认为,等臂染色体是细胞分裂时着丝粒横裂所致,横裂后,两长臂和两短臂分别形成等臂染色体进入2个子细胞。等臂染色体实质上是同时带有整臂缺失和整臂重复的染色体。

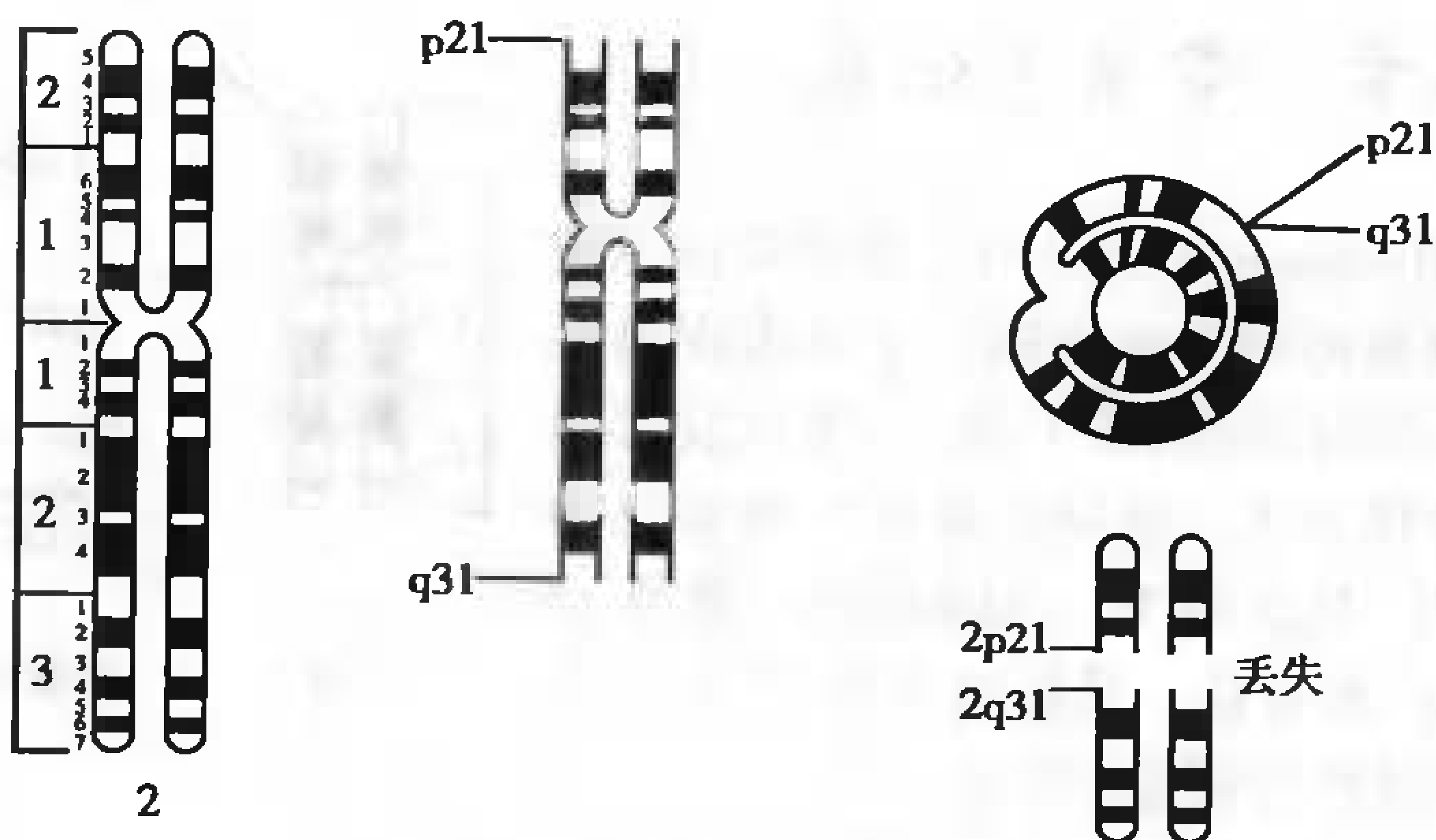


图 7-12 环状染色体

例如:X 染色体长臂构成的等臂染色体(图 7-13),简式为 46,X,i(Xq),详式为 46,X,i(X)(qter→q10::q10→qter)。

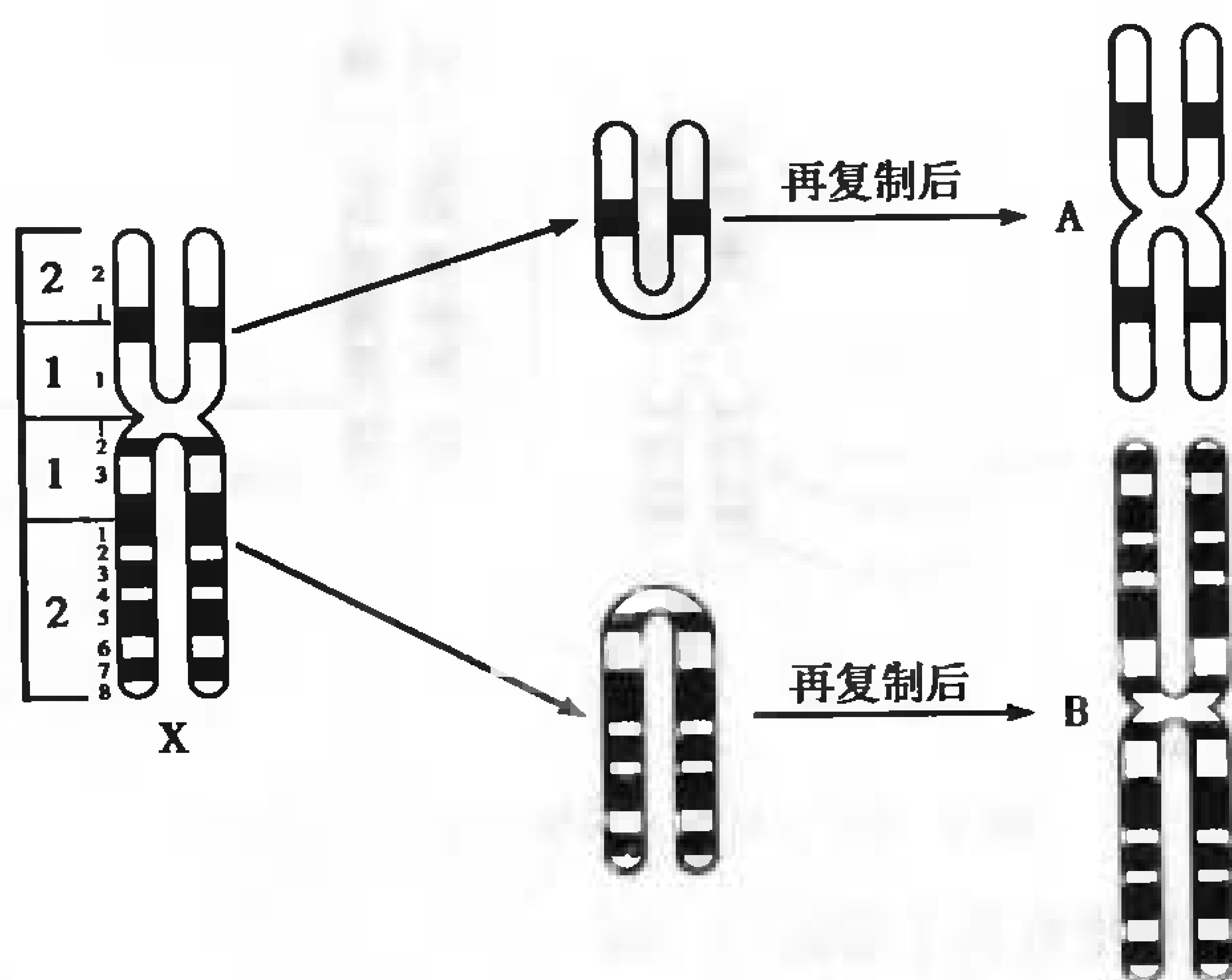


图 7－13　等臂染色体

A:i(Xp)　B:i(Xq)

7.**双着丝粒染色体**(dicentric chromosome)　2 条染色体断裂后，有着丝粒的 2 个片段相接合，形成双着丝粒染色体，2 个无着丝粒片段常丢失。

例如：6 号染色体 q22 和 11 号染色体 p15 断裂形成双着丝粒染色体(图 7－14)，简式为 45，XY，dic(6;11)(q22;p15)，详式为 45，XY，dic(6;11)(6pter→6q22::11p15→11qter)。

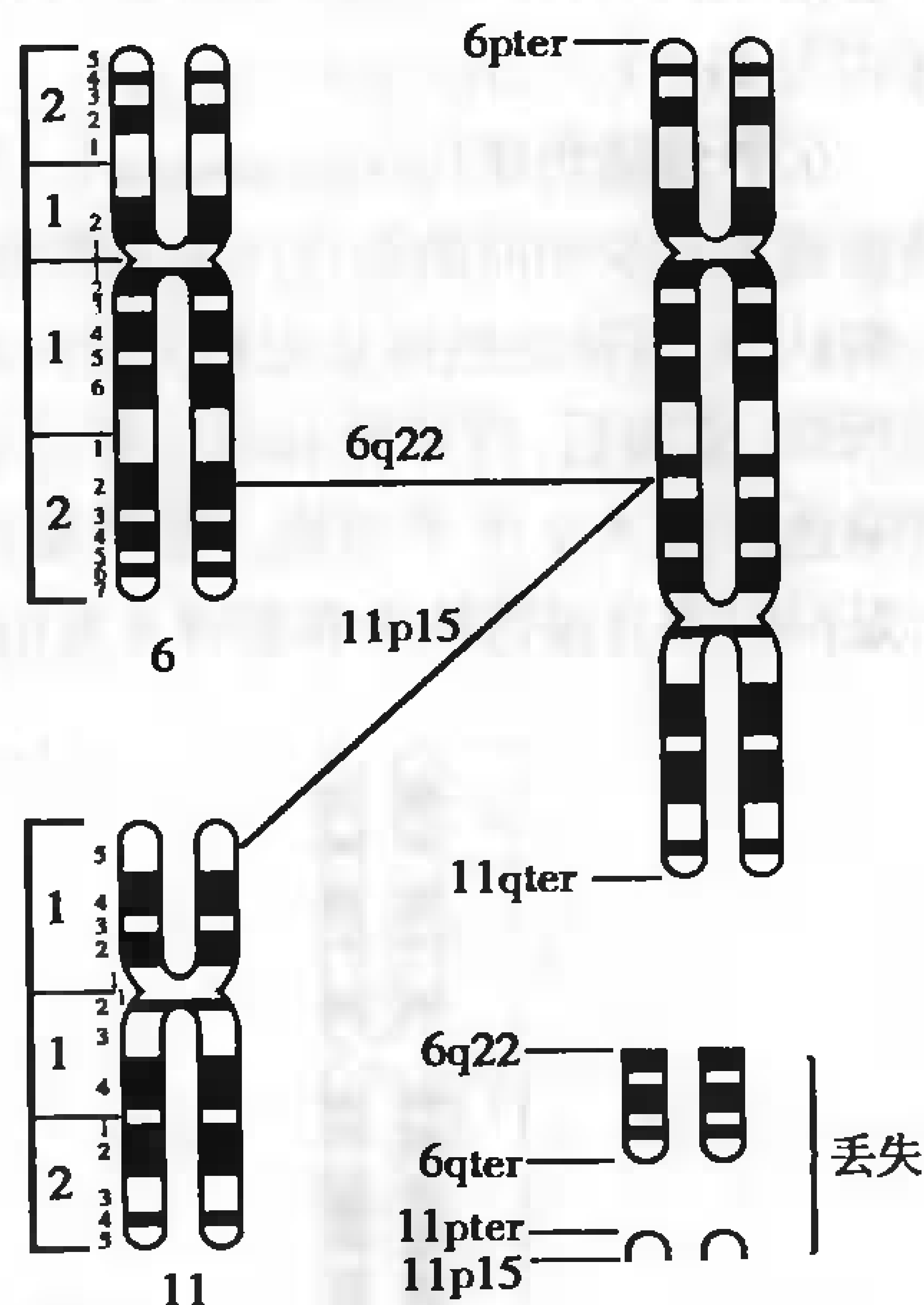


图 7－14　双着丝粒染色体

第二节　常染色体病

常染色体病(autosomal disease)是指常染色体数目异常或结构畸变而引起的疾病。主要表现为先天性缺陷。临床特征包括：智力低下、生长发育延迟及特征性异常体征等。常染色体病中最常见的是 21 三体综合征，其次是 18 三体综合征，偶见 13 三体综合征及 $5p^-$ 综合征。其他如 8 号、22 号三体综合征及部分三体综合征等均罕见。

一、21 三体综合征

21 三体综合征(21 trisomy syndrome)又称先天愚型，1866 年英国医生 Down Langdon 首先描述此病，故又称为 Down 综合征(Down syndrome)。1959 年，法国遗传学家 Lejeune 等证实此病的病因是多了 1 条 G 组染色体，后来用染色体显带技术确定为 21 号染色体，所以

此病称为 21 三体综合征。

21 三体综合征的发病率为 1/800～1/600，无明显性别差异。中度或重度智力低下(mental retardation，MR)是此病最突出的表现，但程度存在差异，智商通常为 25～50，高于 50 的很少。患者有特殊的面部特征(图 7－15)：头小面圆，枕部扁平；睑裂狭小上斜，眼距宽，内眦赘皮，常有斜视；鼻扁平；嘴小唇厚，舌大常伸出口外，流涎；耳小，耳位低；身材矮小；指短，小指内弯且为单一指褶纹；掌纹中三叉点 t 移向掌心(t′)，atd 角增大；通贯手；第 1、2 趾间距宽(草鞋趾)，拇趾球区胫侧弓形纹；肌张力低下，常有腹直肌分离或脐疝；先天性心脏病；男性常有隐睾而没有生育能力，女性通常无月经，生育力下降；先天性免疫力低下，易患呼吸道感染，白血病发病率明显增高。患者寿命短，约 50% 在 5 岁前死亡，平均寿命 16.2 岁。根据患者核型组成的不同，21 三体综合征分为 3 类。



图 7－15 21 三体综合征患者

(一)21 三体型

21 三体型又称为游离型，核型为 47，XX(XY)，＋21，此型约占 95%。患者全身体细胞均比正常人多 1 条 21 号染色体(图 7－16)，临床表现典型。多出的 21 号染色体源于减数分裂不分离。约 95% 发生在母亲的生殖细胞，主要是减数分裂Ⅰ不分离，仅 5% 发生在父亲的生殖细胞。母亲生育年龄是影响发病率的重要因素。母亲大于 35 岁时。发病率明显增高，35 岁时为 1/300，40 岁时为 1/100，45 岁后则升至 1/50(图 7－17)。

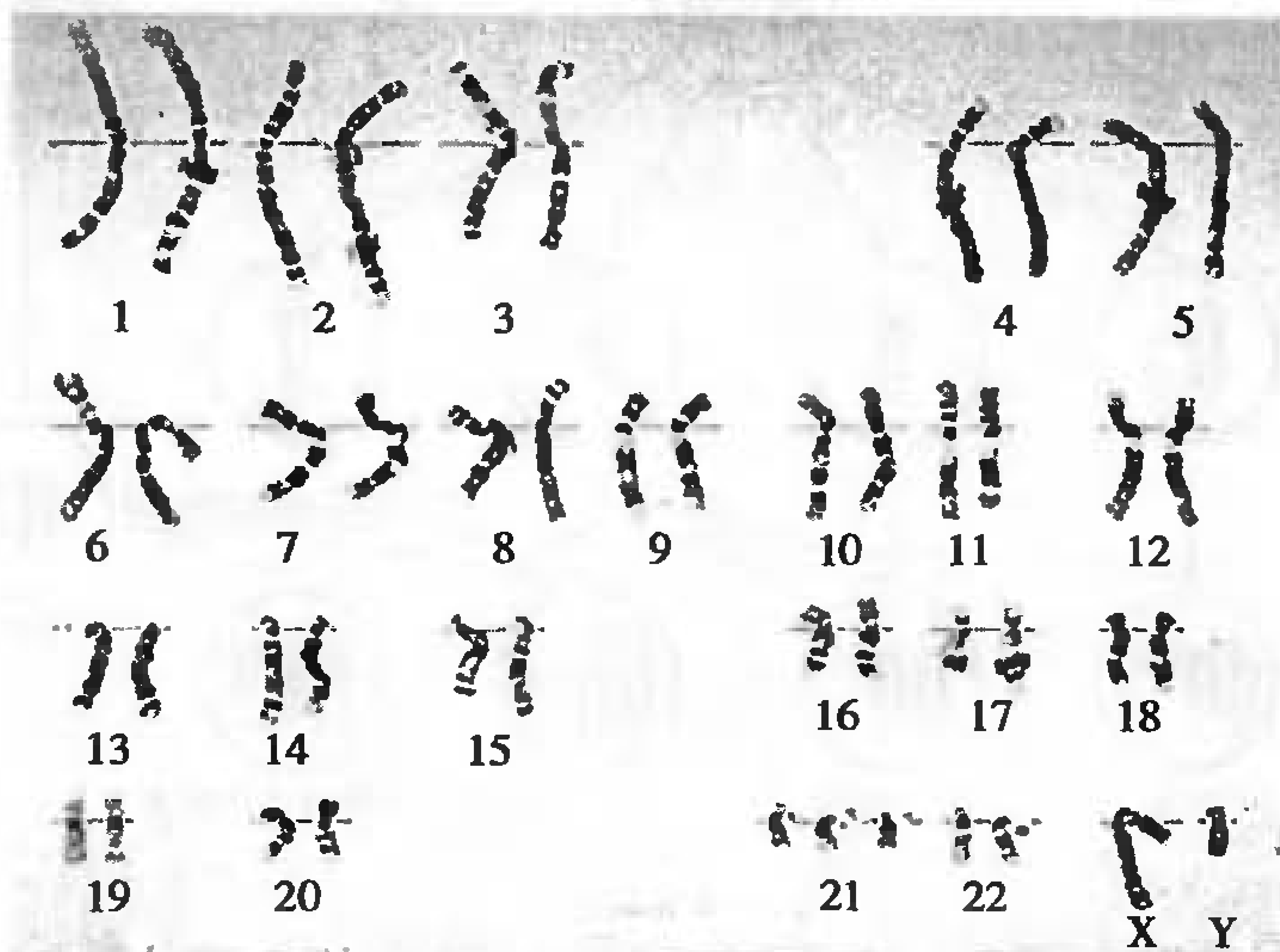


图 7－16 21 三体综合征核型

(二)嵌合型

核型为 47，XX(XY)，＋21/46，XX(XY)，此型占 1%～2%。患者病情取决于异常细胞系所占的比例和它们在体内的分布差异，总的来说，病情比 21 三体型轻，如果异常细胞很少(＜9%)，则表型与正常人无异。

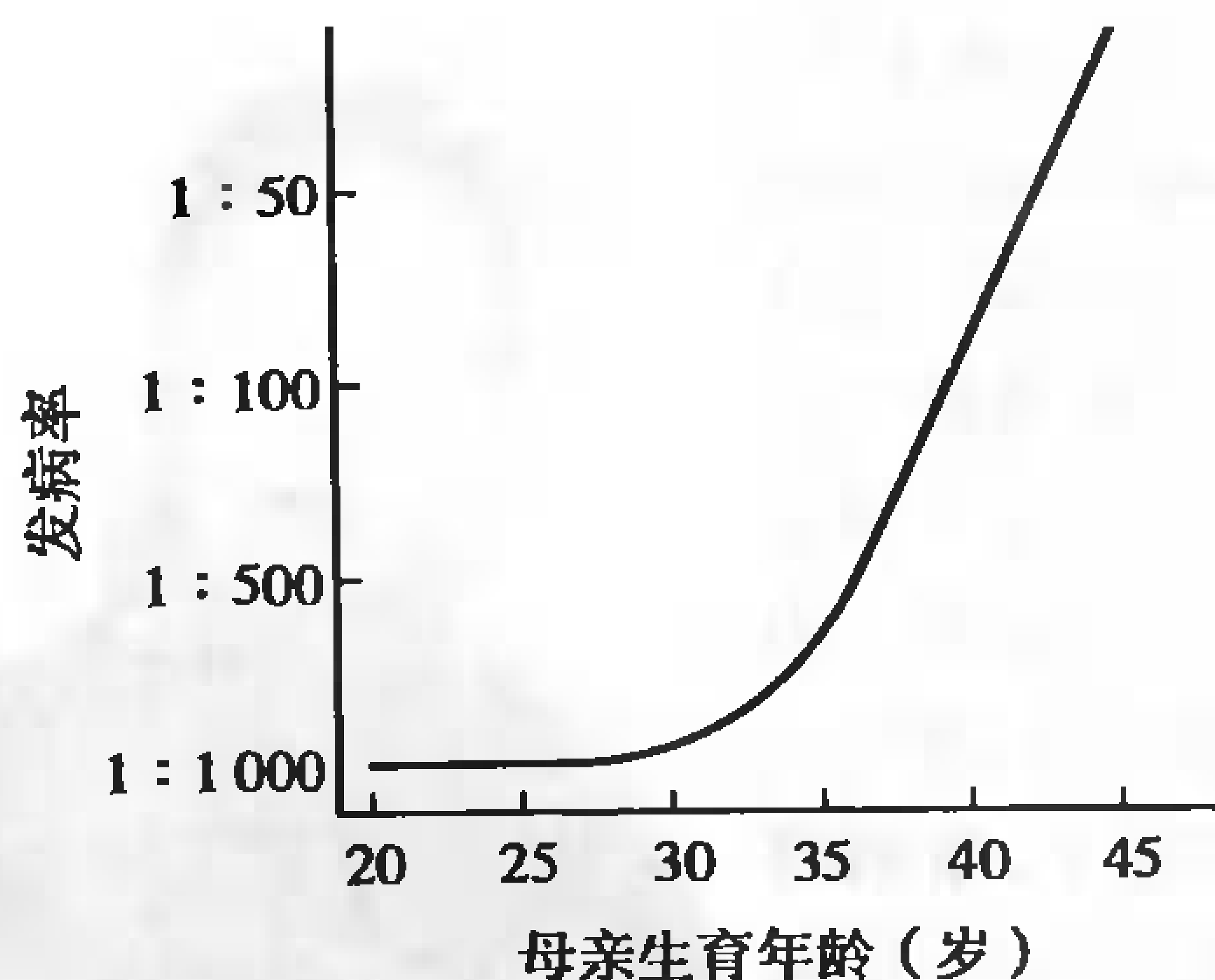


图 7－17　母亲生育年龄与21 三体综合征的发病率

（三）易位型

此型占 3%～4%，患者核型中有 1 条 D 组或 G 组染色体与 21 号染色体经罗伯逊易位形成的易位染色体。因此，此型核型简称为 Dq21q、Gq21q。患者虽然只有 46 条染色体，但因核型中 2 条正常 21 号染色体外多了 1 条易位染色体，而易位染色体含有 21 号染色体长臂绝大部分结构，因此，表现出典型的先天愚型特征，即这是一种假二倍体。

在两类核型中，Dq21q 多于 Gq21q，其中以 14q21q 易位最常见，其核型为：46，XX（XY），der（14；21）（q10；q10），+21。易位染色体约半数是双亲生殖细胞减数分裂前新发生的罗伯逊易位，另一半是表型正常的罗伯逊易位携带者的亲代（主要是母亲）遗传而来。易位携带者的核型为 45，XX（XY），der（14；21）（q10；q10），经减数分裂可形成 4 种配子，若与正常个体婚配，子代有 4 种核型（图 7－18）：正常二倍体、与亲代相同的易位携带者、易位型先天愚型和 21 单体型。

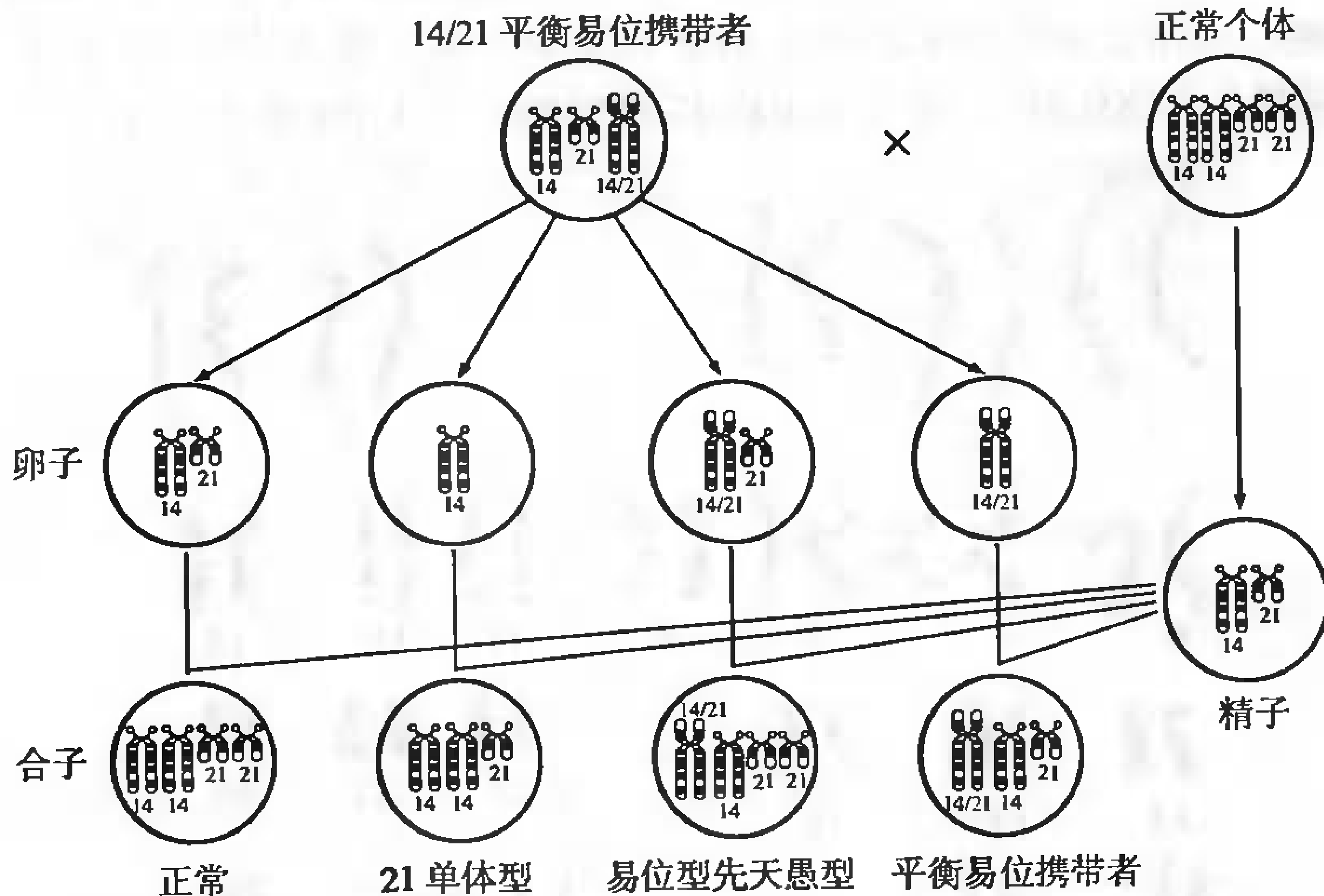


图 7－18　14/21 易位携带者的遗传效应

如果双亲之一是核型为 45，XX（XY），der（21；21）（q10；q10）的 21q21q 罗伯逊易位携带者，与正常个体婚配，子代 1/2 是 21 单体型，核型为：45，XX（XY），－21；1/2 是易位型先天愚型，核型为：46，XX（XY），der（21；21）（q10；q10），+21。21 单体型在胚胎期均流产，出生的都将是易位型先天愚型患者，即活婴 100%受累。因此，21q21q 平衡易位携带者不宜

生育。

与21三体型不同,易位型患者的父母为年轻夫妇。

二、18三体综合征

18三体综合征(18 trisomy syndrome)又称为Edwards综合征(Edwards syndrome),1960年Edwards首先描述本病,1964年,Yunis用放射自显影方法证实多出的一条染色体为18号染色体。新生儿发病率为1/5 000~1/4 000,男女发病率之比为1:4。

本病的临床特征是生命力严重低下,多发畸形,生长和智能发育迟缓。临床表现主要有:枕部凸出;眼裂小,眼球小,内眦赘皮;鼻梁细长,嘴小,耳位低,耳廓畸形;全身骨骼肌发育异常,胸骨短,骨盆狭窄;第1、2、5指压在第3、4指上,互相叠盖,呈特殊的握拳姿势(图7-19);足跟后凸,足掌中部凸出呈"摇椅"状;95%的患儿有先天性心脏病。本病因有多种严重畸形,故大多在胚胎期流产,特别是男胎,这可能与男性发病率低有关。出生儿平均存活71天,个别个体可存活数年。

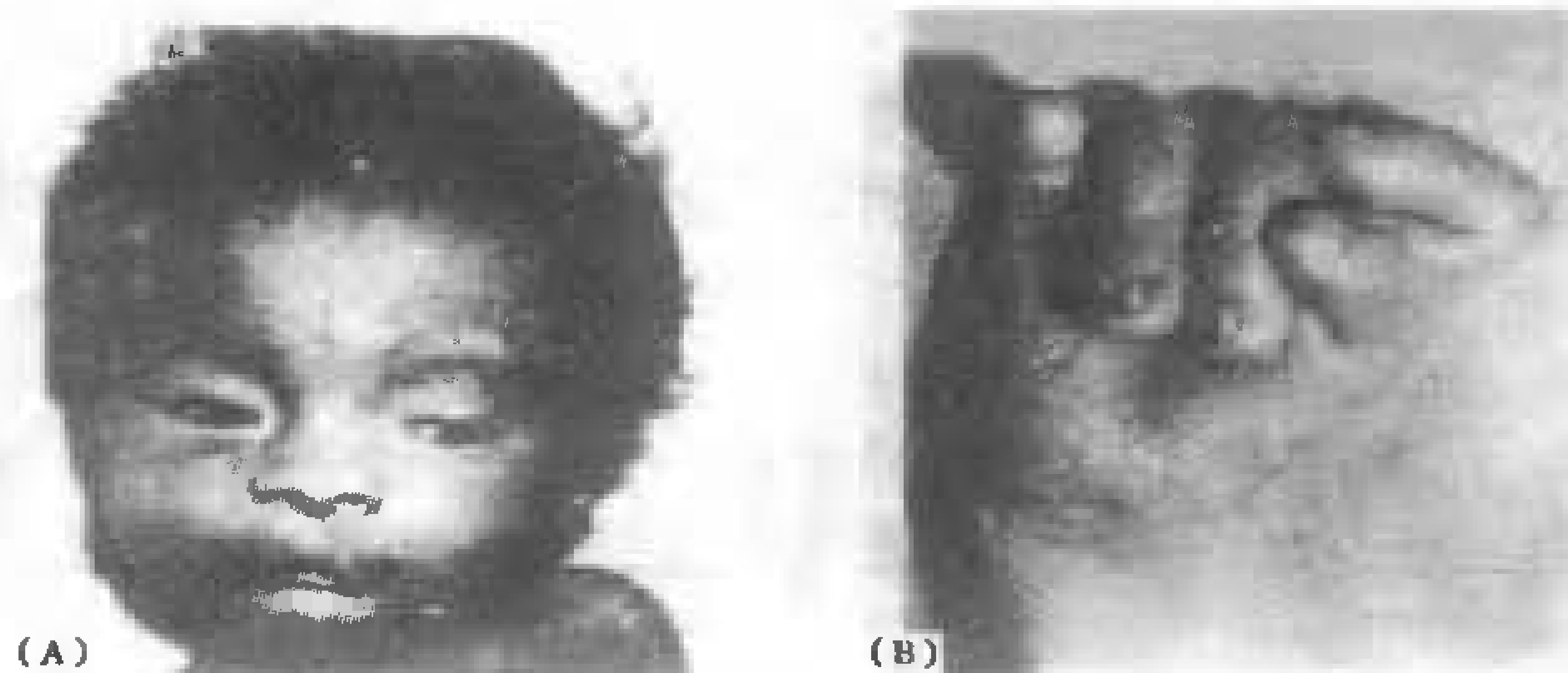


图7-19 18三体综合征患者
A:面部特征 B:特殊的握拳姿势

80%患儿的核型为18三体型,即47,XY(XX),+18(图7-20),临床表现典型。此型多由母亲卵母细胞减数分裂时18号染色体不分离所致,且与母亲的生育年龄有关。少数患者的核型为嵌合型,即47,XY(XX),+18/46,XY(XX);临床表现相对较轻,存活时间稍长,此型与母亲的生育年龄无关。极少数为易位型核型。

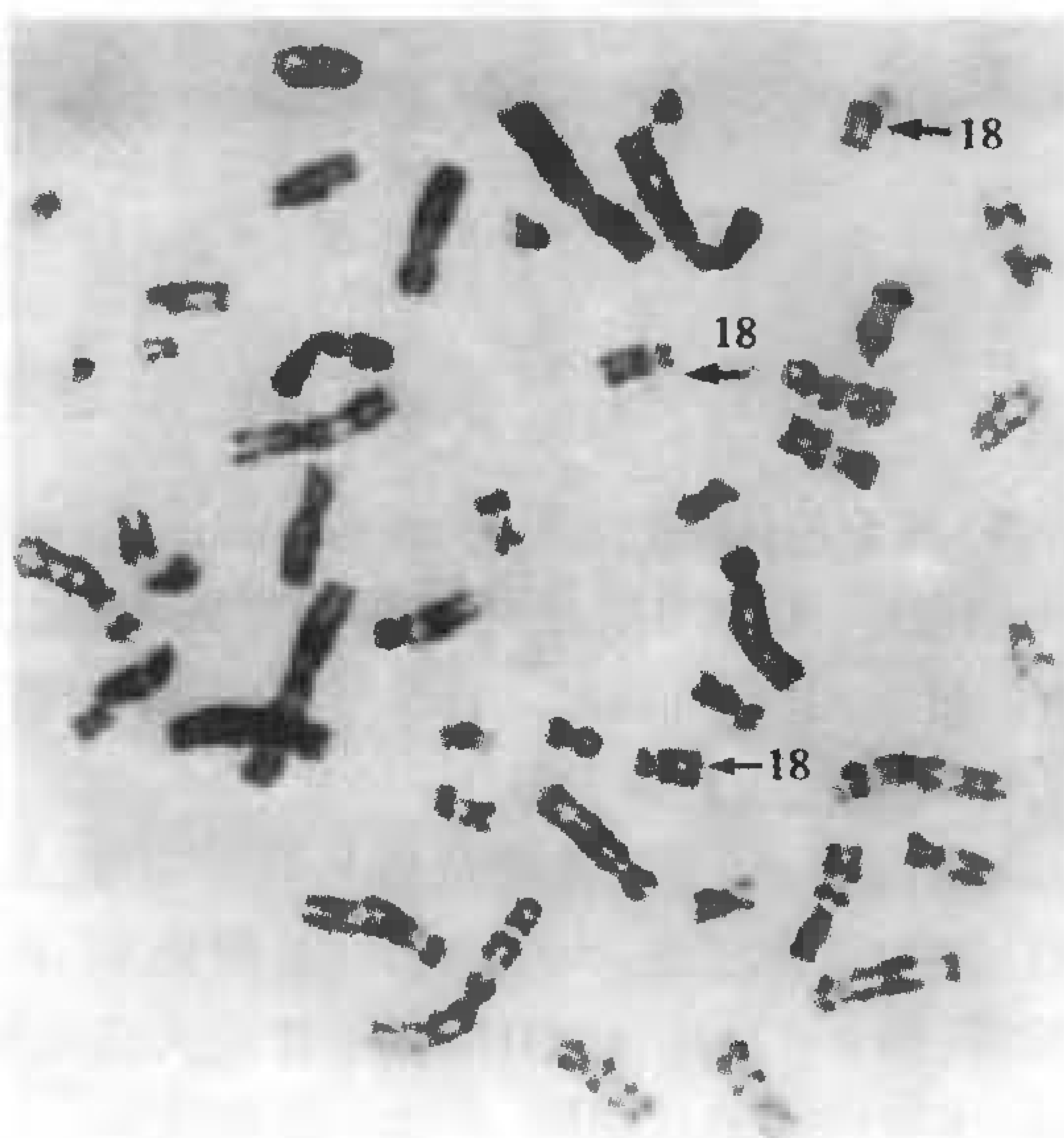


图7-20 18三体综合征核型

三、13三体综合征

13三体综合征(13 trisomy syndrome)又称为Patau综合征(Patau syndrome),1960年,Patau首次报道本病。新生儿发病率约为1/7 000,女性明显多于男性。

此病患儿的临床表现比21三体综合


图 7-21　13 三体综合征患者

征和 18 三体综合征更严重。主要临床表现有：发育迟缓、智能低下；中枢神经系统发育缺陷：前脑皮质缺如，小脑小，无嗅脑；头小，眼小，颌小；鼻大扁平，唇裂，或伴腭裂；耳廓畸形，耳低位；男性阴囊畸形或隐睾，女性阴蒂肥大、双阴道、双角子宫等；内脏多发性畸形，心脏发育异常，多囊肾，无脾等(图 7-21)。99%的 13 三体型胚胎流产，出生儿中，约 45%生后 1 个月内死亡，90%生后 6 个月内死亡，存活至 3 岁者少于 5%。

80%患儿核型为 13 三体型，即 47,XX(XY),+13(图 7-22)，此型发病率与母亲的生育年龄有关。10%～15%为易位型核型，多为 46,XX(XY),+13,der(13;14)(q10;q10)。5%为嵌合型核型，即 47,XX(XY),+13/46,XX(XY)，此型临床表现相对较轻。

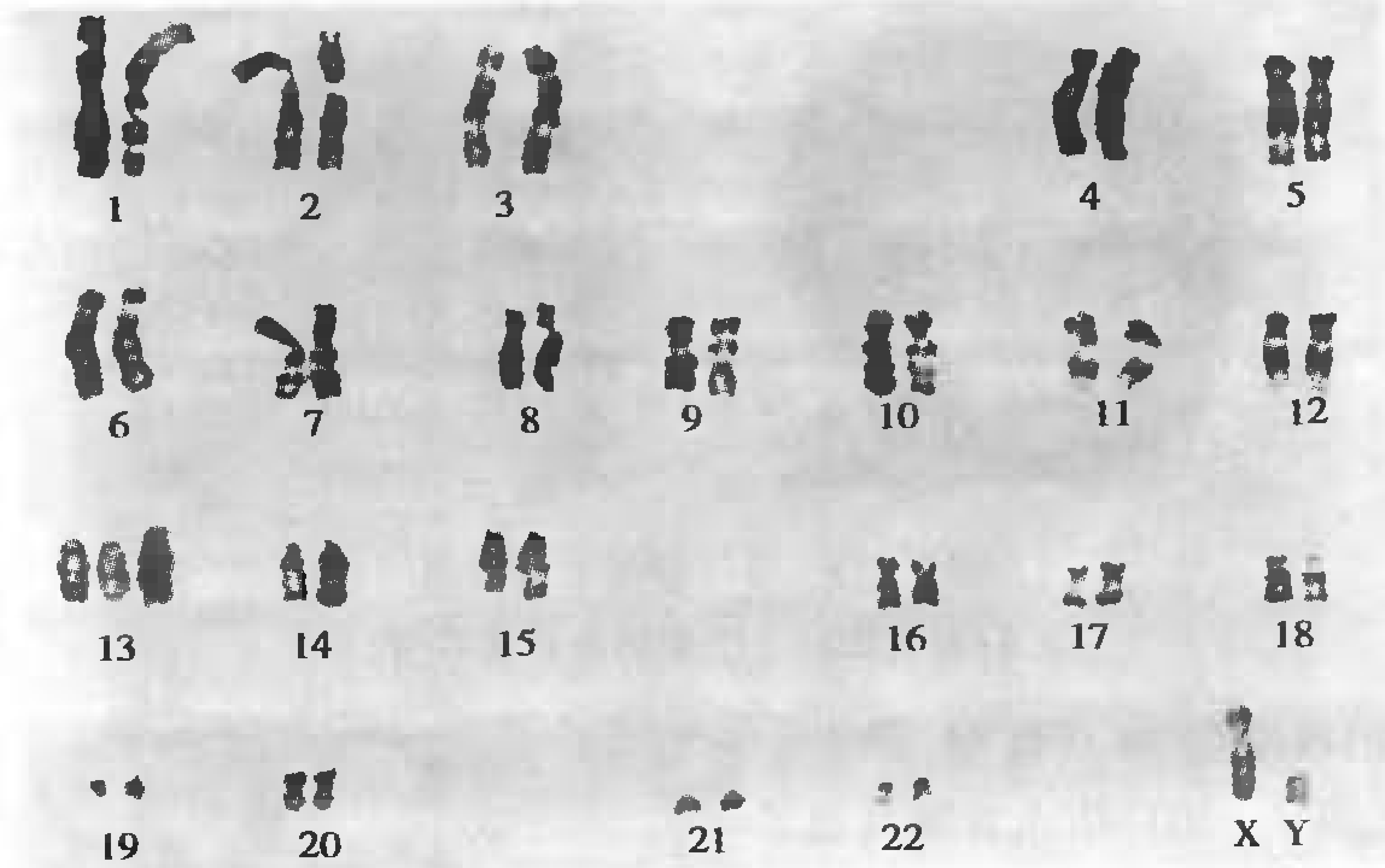


图 7-22　13 三体综合征核型

四、5p 部分单体综合征

5p 部分单体综合征(partial monosomy 5p syndrome)为最常见的缺失综合征，患儿有特征性的猫叫样哭声，故又称为猫叫综合征(Cri-du-chat syndrome)。1963 年，Lejeune 首先报道此病。发病率约为 1/50 000，女性多于男性。

本病最具特征的表现是患儿哭声尖细，似猫叫；其他临床表现有：智力发育迟缓，头小，满月脸，眼距宽，内眦赘皮，下颌小且后缩等(图 7-23)。约 50%患者有先天性心脏病。患儿死亡率较低，多数可活至儿童期，少数可活到成年。


图 7-23　$5p^-$综合征患者

不同患者 5p 缺失($5p^-$)片段大小不同，但都包含 5p15 区域，核型为 46,XX(XY),del(5)(p15)(图 7-24)。这说明患儿的临床

表现是5p15的特异性缺失引起的。

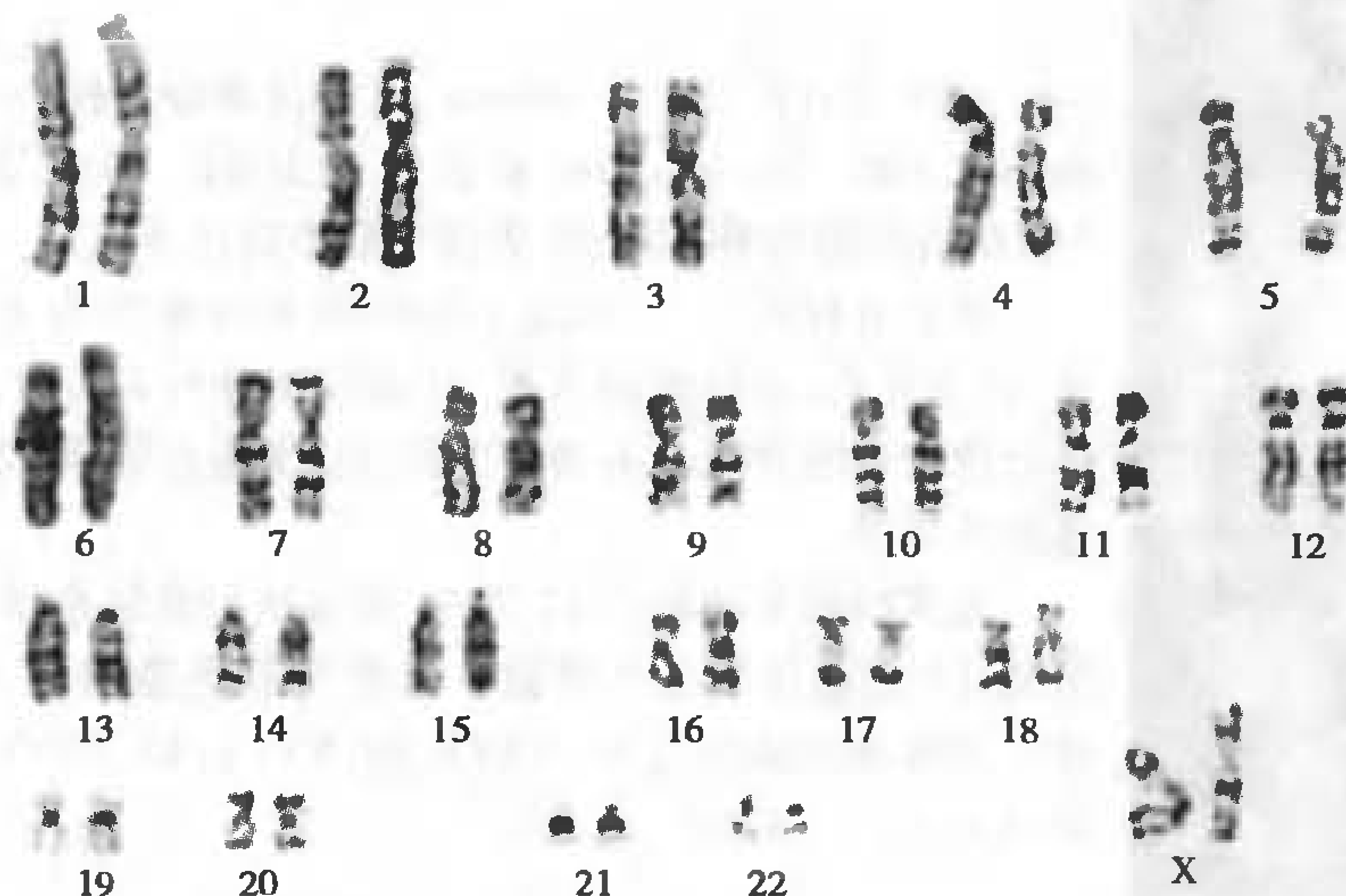


图7-24　5p⁻综合征核型

第三节　性染色体病

性染色体病(sex chromosomal disease)是X和Y染色体数目异常或结构畸变引起的疾病,主要特征是性发育不全或两性畸形。

一、Klinefelter综合征

Klinefelter综合征(Klinefelter syndrome)又称为先天性睾丸发育不全或XXY综合征。1942年Klinefelter首先报道此病。男性发病率约为1/850,占男性不育的1/10。

此病以睾丸发育障碍和不育为主要特征。患者外观男性;身材瘦高,四肢较长;青春期后第二性征发育差,呈现女性化表现:体毛稀少,阴毛少且呈女性分布,无胡须,喉结不显,皮下脂肪丰富、皮肤细嫩,部分患者乳房发育(图7-25);睾丸小(直径小于2 cm)而质硬,曲细精管萎缩或呈玻璃样变,不能产生精子而不育;部分患者轻度智能障碍,或有精神异常及易患精神分裂症倾向。由于此病青春期前临床表现不明显,因而儿童期诊断较难,如果儿童期发现睾丸、阴茎特别小,应考虑此病,可进行性染色质检查或染色体核型分析以便早期诊断。

绝大多数患者核型为47,XXY(图7-26),其中60%是母亲卵母细胞减数分裂不分离所致,患者X染色质、Y染色质均呈阳性。15%患者为嵌合型核型,包括:46,XY/47,XXY;46,XY/48,XXXY。如果正常细胞所占比例较大,一侧睾丸可发育良好并有生育能力。少数患者核型为48,XXXY或49,XXXXY。一般来说,核型中X染色体越多,病情越重。



图 7-25 Klinefelter 综合征患者(示男性乳房发育)

二、XYY 综合征

XYY 综合征(XYY syndrome)又称超雄综合征(supermale syndrome),1961 年,Sandburg 首次报道此病。男性发病率约为 1/900,在监狱和精神病院,男性发病率高达 3%。

患者身材高大,2 米以上男性的发病率约为 10%;暴躁粗鲁,行为过火,自我克制力差,常有攻击性行为、反社会行为;大部分患者性发育正常并有生育能力,少数有隐睾、睾丸发育差、尿道下裂等。

大多数患者核型为 47,XYY,异常核型是父亲精原细胞减数分裂Ⅱ Y 染色体不分离所致。患者 X 染色质阴性、Y 染色质阳性。少数患者核型为 48,XXYY;48,XYYY;49,XYYYY;47,XYY/46,XY;45,X/49,XYYYY 等。

三、Turner 综合征

Turner 综合征(Turner syndrome)又称为先天性性腺发育不全征或先天性卵巢发育不全综合征。1938 年,Turner 首次描述此病。新生女婴发病率约为 1/2 500,但自发流产胚胎中,Turner 综合征比例高达 7.5%。

患者外观女性,身材矮小,成年身高 120 ~ 140 cm。卵巢呈条索状,无卵泡;子宫小,原发性闭经;阴毛、腋毛稀少,甚至缺如,外生殖器呈幼稚型;盾状胸,乳房发育差,乳间距宽;后发际很低,蹼颈,肘外翻;婴儿期足背淋巴水肿(图 7-27)等。患者大多能存活至成年。确诊患者青春期后用雌激素治疗,可促进第二性征发育。

60%患者的核型为 45,X(图 7-28),其中 80%起源于父亲精原细胞减数分裂Ⅱ不分离。X 染色质、Y 染色质均呈阴性,临床表现最典型。

除 45,X 外,Turner 综合征核型还有多种,包括:嵌合型,45,X/46,XX 或 45,X/47,XXX 或 45,X/46,XX/47,XXX;X 染色体等臂,46,X,i(Xq)或 46,X,i(Xp);X 染色体缺失,46,XXp- 或 46,XXq-;X 环状染色体,46,X,r(X)等。

四、多 X 综合征

多 X 综合征(poly-X syndrome)又称超雌综合征(superfemale syndrome),1959 年,Jacobs 首先报道此病,此病发病率约为 1/1 000。

患者外观如正常女性,能生育;多数患者智能低下,学习能力差,人际关系不良,并有易患精神病倾向;少数患者卵巢功能异常,月经失调,乳腺发育不良,不育。

患者核型有多种,大多为 47,XXX,主要是卵子发生时减数分裂Ⅰ不分离所致,X 染色质数为 2;少数为嵌合型,46,XX/47,XXX;其他尚有 48,XXXX;49,XXXXX。X 染色体越多,病情越重。

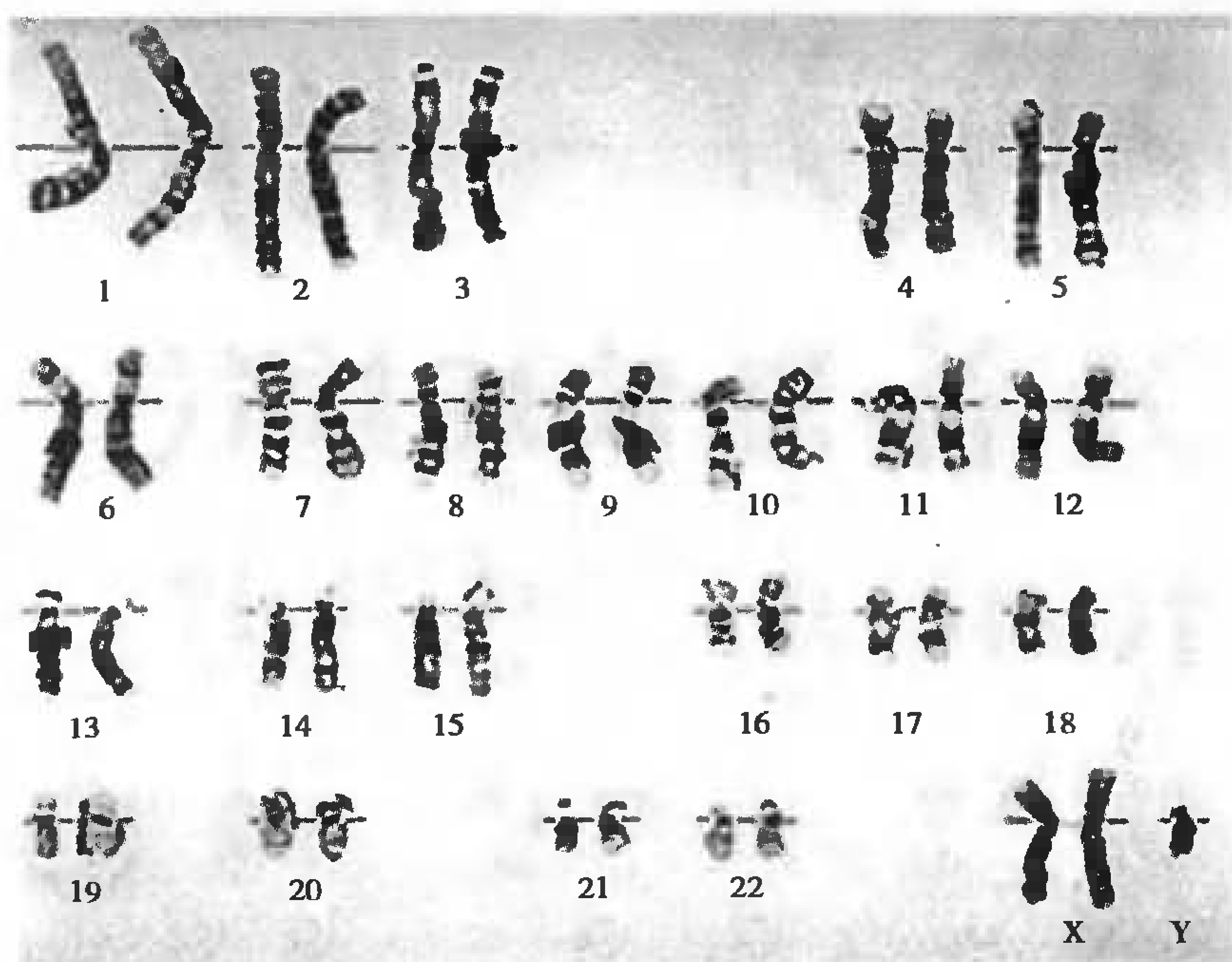


图 7-26 Klinefelter 综合征核型

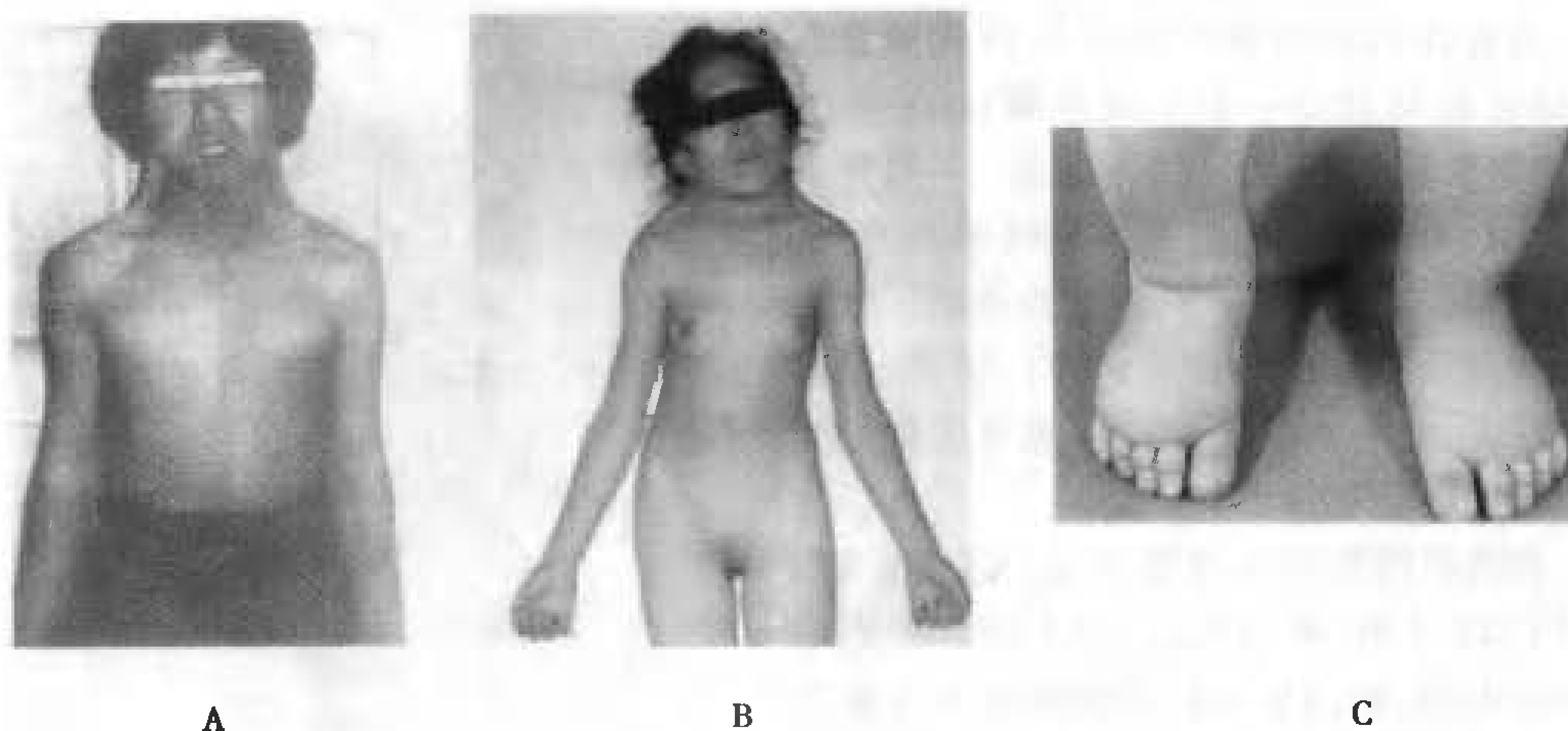


图 7-27 Turner 综合征患者

A.蹼颈 B.乳间距宽、肘外翻 C.足背淋巴水肿

五、两性畸形

两性畸形(hermaphroditism)是指患者的性腺、外生殖器、副性征具有不同程度的两性特征。根据患者体内是否有两性性腺,两性畸形分为真两性畸形和假两性畸形。

(一)真两性畸形

真两性畸形(true hermaphroditism)患者体内兼有睾丸和卵巢,外生殖器不同程度地介

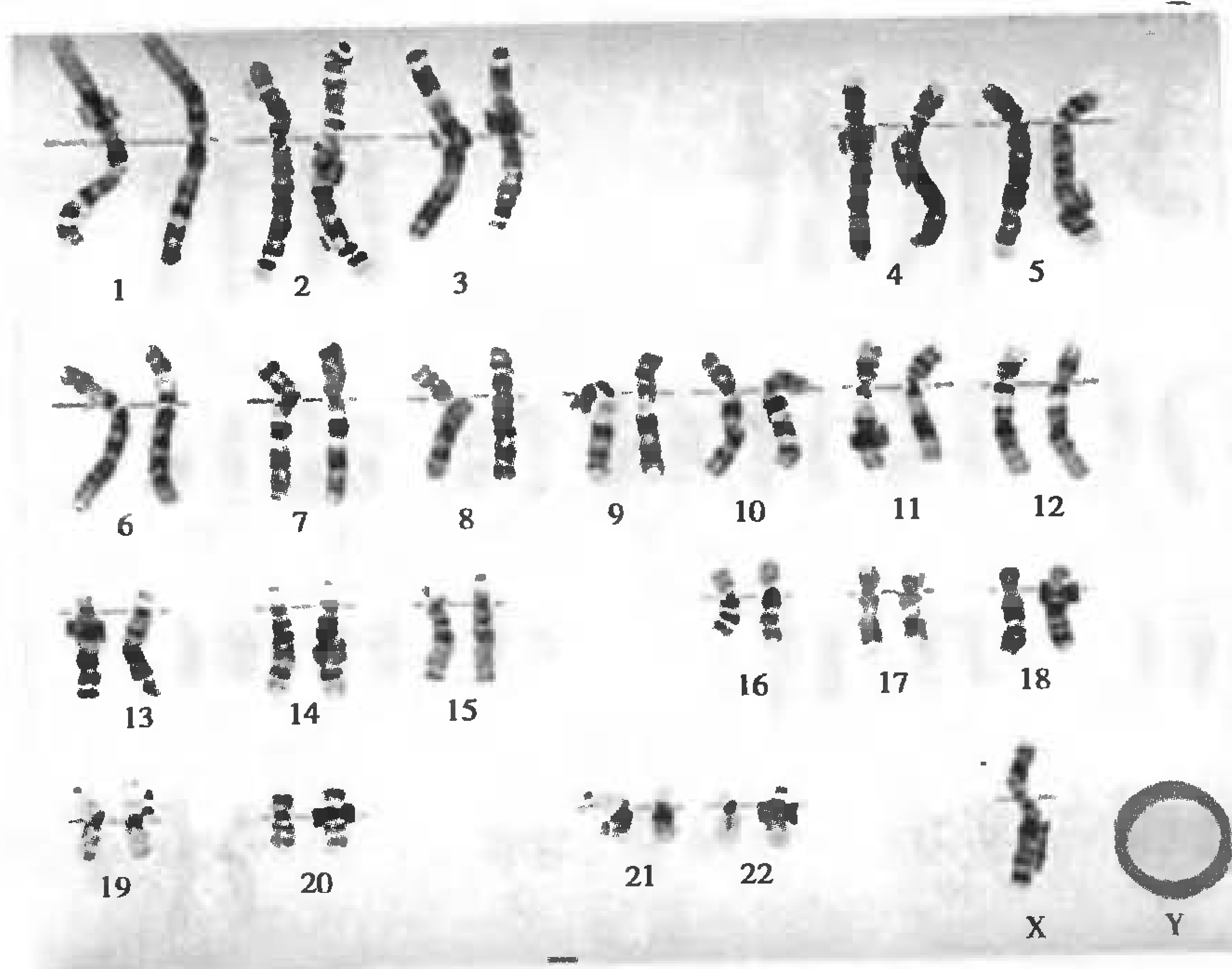


图 7－28　Turner 综合征核型

于两性之间，外观可为男性，也可为女性。

患者体内的两种性腺可以彼此单独存在，也可能结合在一起形成卵睾（ovotestis）。约 40% 患者的性腺一侧为卵巢，一侧为睾丸；另有 40% 患者一侧为卵巢或睾丸，一侧为卵睾；其他 20% 患者两侧均为卵睾。发育不全或部分正常的输卵管、子宫、附睾、精囊同时存在，外生殖器难辨，2/3 患者男性外生殖器较明显（图 7－29）。



图 7－29　真两性畸形患者的外生殖器

真两性畸形的核型包括 46，XX；46，XY；46，XY/45，X 或 46，XX/47，XXY 等。荧光原位杂交表明，46，XX 型真两性畸形患者常染色体或 X 染色体上有易位的 SRY 基因；一般认为，46，XY 型真两性畸形患者体内有部分 46，XX 或 45，X 核型的细胞。

（二）假两性畸形

假两性畸形（pseudohermaphroditism）患者只有一种性腺，但第二性征和外生殖器兼有两性特征或倾向于与性腺相反的性别。

核型为 46，XY、性腺为睾丸的为男性假两性畸形。患者 X 染色质阴性，Y 染色质阳性，体内雄激素水平正常，但由于基因突变致靶细胞雄激素受体缺陷，靶细胞对雄激素的反应不敏感或无反应，因而外生殖器和副性征趋向女性。此型也称睾丸女性化综合征或

雄激素不敏感综合征。

核型为46,XX、性腺为卵巢的为女性假两性畸形。患者X染色质阳性,Y染色质阴性。此型的发生多数是由于患者21-羟化酶遗传性(AR)缺乏,使氢化皮质酮不足,从而使垂体的促肾上腺皮质激素(ACTH)分泌增多,肾上腺皮质过度增生,雄激素合成过多所致;也可由于母体肾上腺皮质功能亢进或在孕期为防流产使用了过多的黄体酮,从而导致血液中雄激素水平增高,导致女胎性发育趋向男性。

两性畸形的治疗不同于其他疾病,做出治疗决定要慎重。染色体核型性别并非是患者应达到的最终指标,要充分考虑患者的表型、社会性别、性腺及外生殖器发育情况、年龄、个人要求等方面。如果通过内科或外科治疗,患者作为某一性别的人可以有相对正常的性生活,即使不育,也是最佳选择。除非在婴儿期,一般不主张改变患者的社会性别。隐睾、卵睾易恶变,宜尽早切除。

六、脆性X染色体综合征

X染色体特定部位(Xq27.3)的结构变异,导致易断裂,并可形成以细丝相连的随体样结构。这样的X染色体称为脆性X染色体(fragile X chromosome, Fra X)(图7-30)。脆性X染色体综合征(fragile X chromosome syndrome)是指一种以脆性X染色体为特征的X连锁遗传的智能低下综合征。

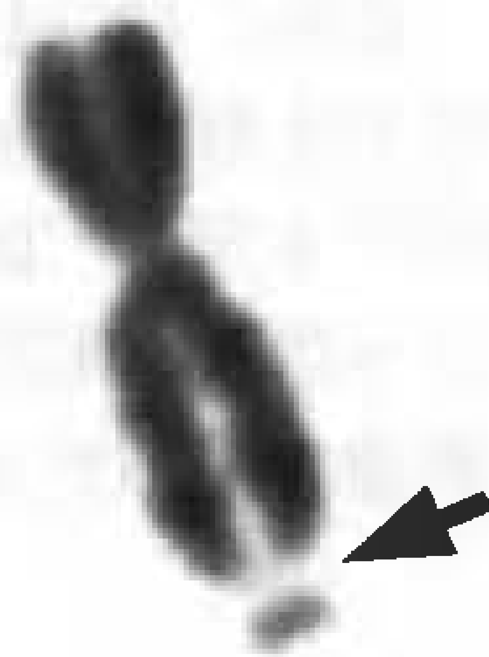


图7-30 脆性X染色体(Fra X)

本病主要是男性发病,占全部男性智力低下的10%~20%。多数男性患者表现为中度到重度的智能低下,智商为20~60;女性多为携带者,由于女性2条X染色体中有1条随机失活,故女性携带者约1/3有轻度智能低下。巨大睾丸是青春期后的典型表现,其他异常包括:特殊的长形面容:前额及下颌凸出,头大、耳大,嘴大唇厚;部分患者有多动症,行为异常,性情孤僻(图7-31)等。



图7-31 脆性X染色体(Fra X)综合征患者
A:患者外貌 B:大睾丸

X染色体的q27.3处存在*FMR*-1基因,该基因5′非翻译区有一段不稳定的(CGG)*n*三核苷酸重复序列,其附近存在CpG序列,又称CpG岛。正常人(CGG)*n*拷贝数6~54,

平均为 30,CpG 岛无甲基化现象;(CGG)*n* 过度扩增和 CpG 岛甲基化在细胞水平表现 Fra X,并导致本病发生。(CGG)*n* 过度扩增主要发生在减数分裂过程中,表现为世代间拷贝数的改变,称为动态突变(参见第二章　遗传的分子基础)。本病携带者(CGG)*n* 拷贝数为 54～200,称前突变(premutation);前突变无临床表现。患者的(CGG)*n* 拷贝数超过 200,称为全突变(full mutation);全突变时出现智能低下及其他脆性 X 染色体综合征的表现,重症脆性 X 染色体综合征患者(CGG)*n* 拷贝数最多可达2 000。

FMR－1 基因编码的 FMR－1 蛋白是一种 RNA 结合蛋白,该蛋白对转录产物的剪切、RNA 的转运、mRNA 的稳定性及翻译有重要的调节作用。FMR－1 蛋白在胎儿中枢神经系统发育过程中许多处于增殖和迁移状态的细胞中有表达。突变,特别是全突变,使 *FMR*－1 基因活性下降,FMR－1 蛋白减少,影响皮质和边缘系统的发育,从而影响依赖于皮质和边缘系统的意向、学习和记忆行为等,表现为智能发育障碍。

第四节　染色体畸变携带者

染色体畸变携带者是指带有染色体结构畸变,但染色体物质的总量基本上仍为二倍体的表型正常的个体,包括易位携带者和倒位携带者两类。染色体畸变携带者虽然表型正常,但可产生部分三体和部分单体的子代,导致不育、流产、死产、新生儿死亡、先天性畸形和智力低下等。在不育与流产夫妇中,染色体畸变携带者占 3%～6%。至今已记载的染色体畸变携带者 16 000 余种,我国已记载 1 200 余种,几乎涉及到每号染色体的每个区带。

一、非同源染色体相互易位携带者

非同源染色体相互易位携带者缺少 2 条正常的非同源染色体,多了 2 条衍生染色体。如夫妇一方为 2/5 染色体平衡易位携带者,核型为 46,XX(XY),t(2;5)(q21;q31),根据配子形成中同源染色体节段相互配对的特性,在减数分裂前期Ⅰ时,易位的染色体和正常染色体联会,配对形成四射体(图 7－32);经过分离、交换、组合,理论上可形成 18 种配子。

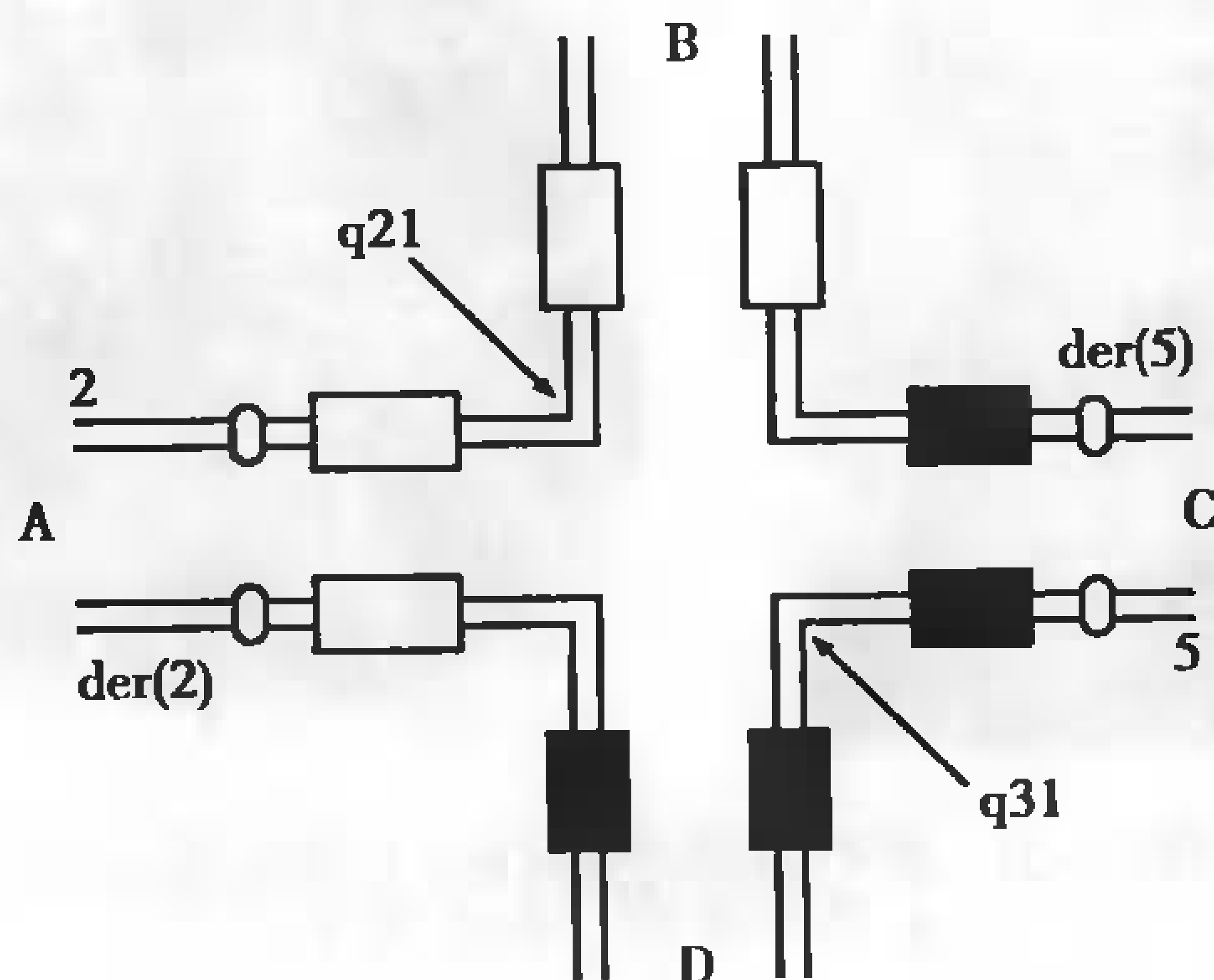


图 7－32　相互易位染色体在减数分裂前期Ⅰ形成四射体图解
A、B、C、D 代表各染色体末端至断裂点的染色体片段

其中,1 种是正常配子,1 种是平衡易位的配子,其余 16 种都是不平衡的;与正常配子受精后,1 种为正常个体,1 种为易位携带者,其余 16 种为单体、部分单体、三体、部分三体,将导致患胎流产、死胎或畸形(表 7 - 2)。

表 7 - 2 相互易位携带者产生的 18 种配子与正常配子受精后的合子类型

分离类型		与正常配子受精后产生的合子类型
对位	AB CD	46,XX/XY
	AD CB	46,XX/XY, -2, -5, +der(2), +der(5),t(2;5)(q21;q31)
邻位-1	AB CB	46,XX/XY, -5, +der(5),t(2;5)(q21;q31)
	AD CD	46,XX/XY, -2, +der(2),t(2;5)(q21;q31)
邻位-2	AB AD	46,XX/XY, -5, +der(2),t(2;5)(q21;q31)
	CB CD	46,XX/XY, -2, +der(5),t(2;5)(q21;q31)
	* AB AB	46,XX/XY, +2, -5
	* CD CD	46,XX/XY, -2, +5
	* CB CB	46,XX/XY, -2, -5, +2der(5),t(2;5)(q21;q31)
	* AD AD	46,XX/XY, -2, -5, +2der(2),t(2;5)(q21;q31)
3:1 分离	AB CB CD	47,XX/XY, +der(5),t(2;5)(q21;q31)
	AD	45,XX/XY, -2, -5, +der(2),t(2;5)(q21;q31)
	CB CD AD	47,XX/XY, -2, +der(2), +der(5),t(2;5)(q21;q31)
	AB	45,XX/XY, -5
	CD AD AB	47,XX/XY, +der(2),t(2;5)(q21;q31)
	CB	45,XX/XY, -2, -5, +der(5),t(2;5)(q21;q31)
	AD AB CB	47,XX/XY, -5, +der(2), +der(5),t(2;5)(q21;q31)
	CD	45,XX/XY, -2

* 表示在着丝粒与互换点之间发生交换所形成的配子

二、倒位携带者

臂间倒位携带者表型正常,但在配子形成的减数分裂过程中,根据配子形成时同源染色体的同源节段相互配对规律,臂间倒位的染色体形成特有的倒位圈(图 7 - 33);经过在倒位圈内的交换,理论上可形成 4 种配子;4 种配子与正常配子受精形成的 4 种受精卵,一种正常,一种为倒位携带者,另二种均为部分重复和部分缺失所致的部分单体和部分三体患者(图 7 - 34)。一般说来,倒位片段越短,重复和缺失的部分则越长,受精卵正常发育的可能越小,临床上表现为婚后不育、流产、死产的比例越高;反之,倒位片段越长,重复和缺失的部分则越短,受精卵正常发育的可能性越大,生出畸形儿的危险性相对较少。因此,携带者的检出及携带者妊娠期的产前诊断,对防止患儿出生意义重大。

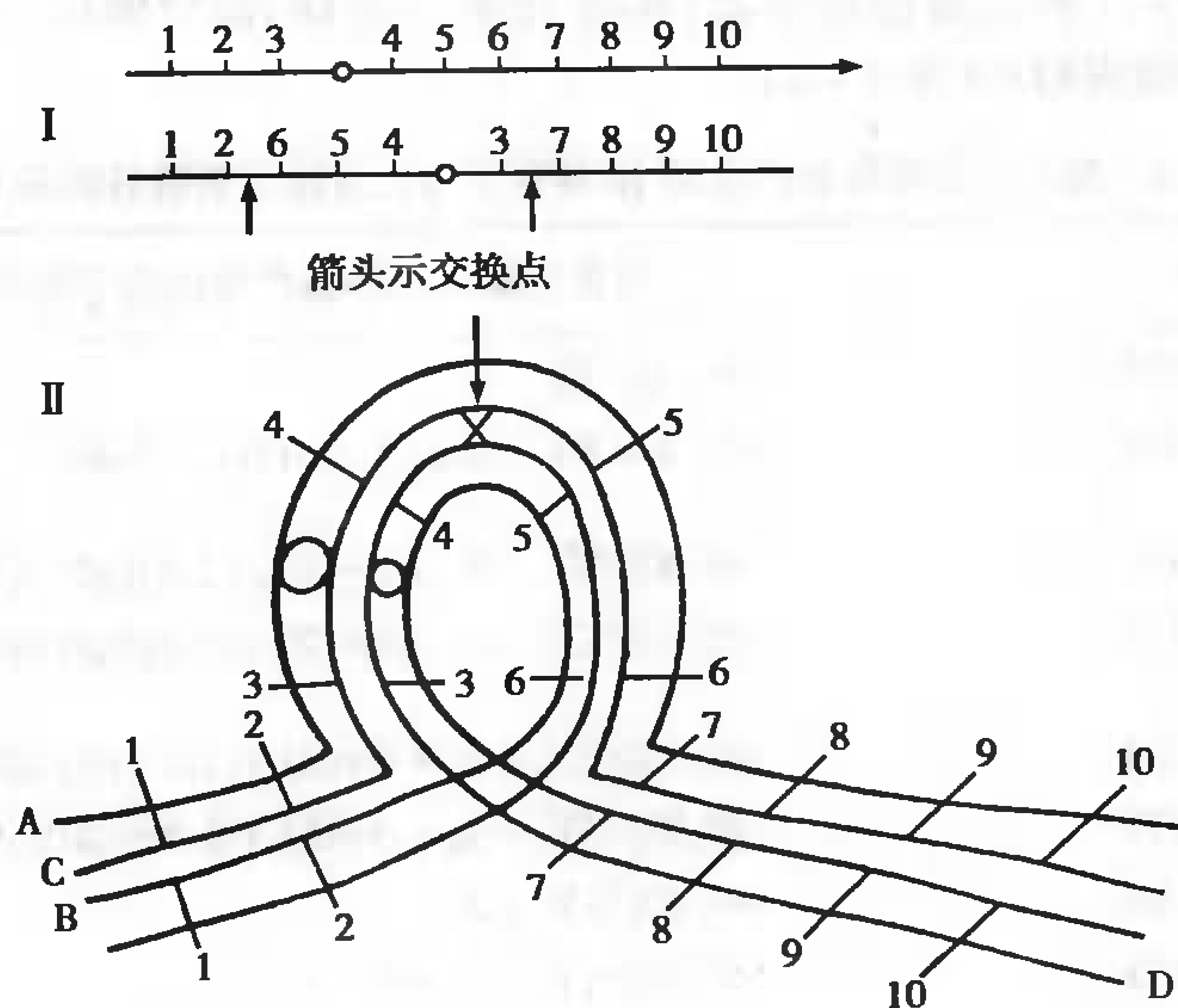


图 7-33　臂间倒位染色体减数分裂前期Ⅰ形成倒位圈图解

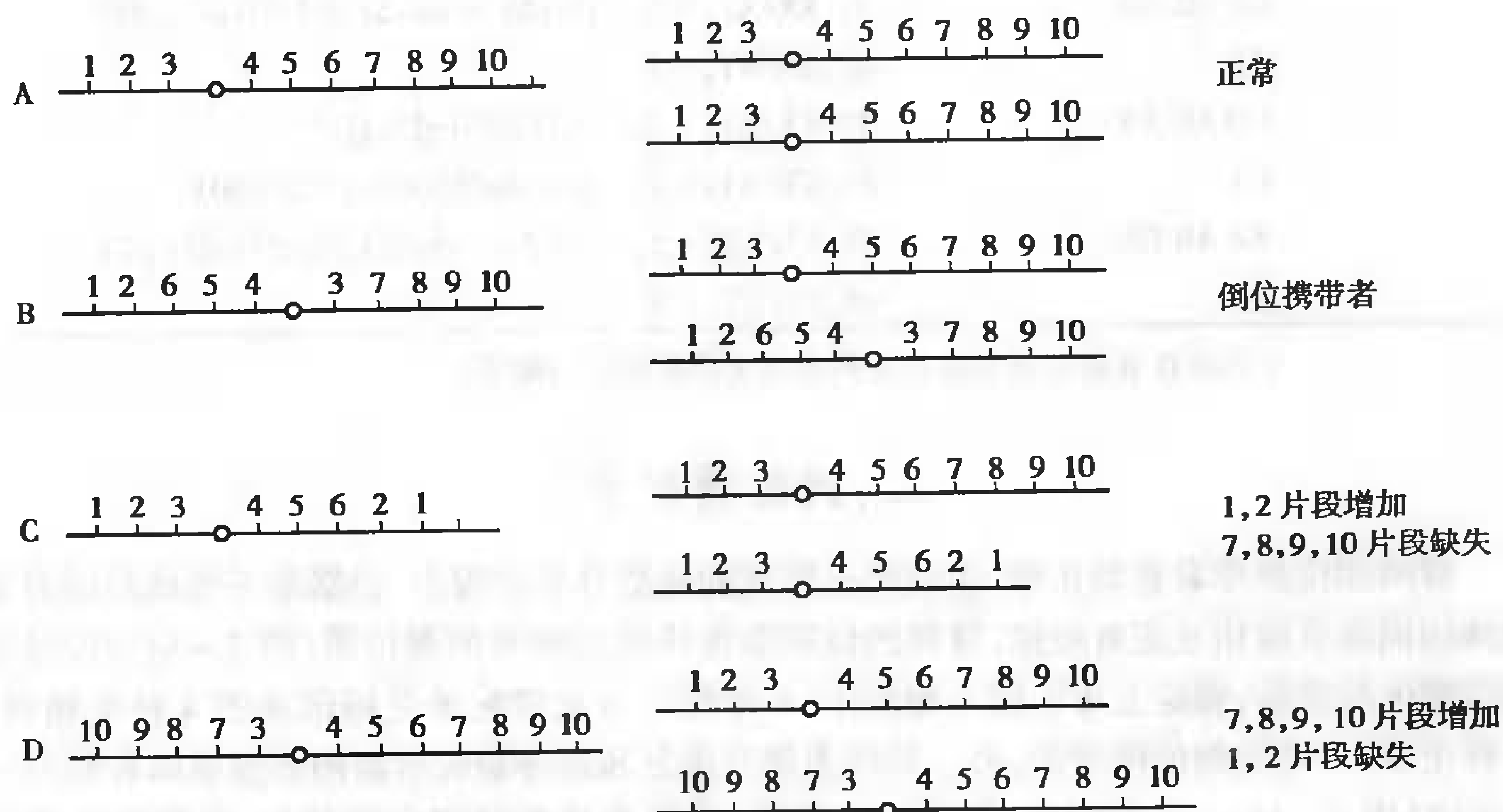


图 7-34　臂间倒位携带者的遗传效应

A:正常配子与正常配子结合后的合子类型

B、C、D:臂间倒位携带者产生的配子与正常配子结合后的合子类型

小 结

染色体病是指染色体数目或结构的畸变引起的疾病,染色体畸变是染色体病形成的根本原因。染色体畸变包括数目异常和结构畸变两类。染色体畸变是在某些化学物质、射线、病毒和机体内在因素作用下产生的。

染色体数目异常包括整倍体异常(三倍体、四倍体和多倍体)、非整倍体异常(超二倍体、亚二倍体和假二倍体)和嵌合体3类,其中较常见的是非整倍体异常是由于细胞分裂机制异常,染色体不分离或染色体丢失所致。

染色体结构畸变类型主要有缺失(中间缺失和末端缺失)、倒位(臂内倒位和臂间倒位)、易位(非相互易位、相互易位和罗伯逊易位)、重复、环状染色体、等臂染色体、双着丝粒染色体等。染色体断裂和(或)异常接合是染色体结构畸变的基础。畸变染色体核型的描述分简式和详式2种,简式对染色体的结构改变只用断裂点表示;详式对染色体的结构改变用重排染色体带的组成来表示。

染色体病分为常染色体病和性染色体病两类。常染色体病是常染色体数目异常或结构畸变引起的疾病,主要表现为先天性缺陷,临床特征包括:智力低下、生长发育延迟及特征性异常体征等。常见的常染色体病有21三体综合征、18三体综合征、13三体综合征及5p-综合征等。

性染色体病是X和Y染色体数目异常或结构畸变引起的疾病。主要特征是性发育不全或两性畸形,常见的性染色体病有Klinefelter综合征、XYY综合征、Turner综合征、多X综合征、脆性X染色体综合征和两性畸形等。

两性畸形是指患者的性腺、外生殖器、副性征具有不同程度的两性特征,分为真两性畸形和假两性畸形两类。真两性畸形患者体内兼有睾丸和卵巢两种性腺组织,其外生殖器不同程度地介于两性之间。假两性畸形患者只有一种性腺,但第二性征和外生殖器兼有两性特征或倾向于与性腺相反的性别。

染色体畸变携带者是指带有染色体结构畸变,但染色体物质的总量基本上仍为二倍体的表型正常的个体,包括易位携带者和倒位携带者两类。染色体畸变携带者虽然表型正常,但可产生部分三体和部分单体的子代,导致不育、流产、死产、新生儿死亡、先天性畸形和智力低下等。

(马志敏 张 励 蔡绍京)

第八章　分子病与先天性代谢病

人类单基因遗传病是基因突变引起的。突变基因编码蛋白质的结构或数量上的变化，将导致机体出现病理变化及功能障碍，这种由于蛋白质分子异常引起的疾病称为分子病（molecular disease）；编码酶蛋白的基因突变将引起酶合成障碍，使机体的代谢过程不能正常进行，最终则导致疾病的发生，这种由于遗传性酶缺陷所致的疾病称为先天性代谢病（inborn errors of metabolism）或遗传性酶病（hereditary enzymopathy）。

第一节　分子病

分子病种类很多，从理论上讲，人体内有多少种蛋白质就可能有多少种分子病。自1949年Pauling等提出分子病的概念以来，已发现这类疾病达2 000多种。本节主要介绍血红蛋白病、血浆蛋白病、受体蛋白病和膜转运载体蛋白病。

一、血红蛋白病

血红蛋白病（hemoglobinopathy）是一类与珠蛋白缺陷有关的疾病。血红蛋白病分为2类：一类为珠蛋白基因异常，导致珠蛋白肽链的结构及功能异常所致的异常血红蛋白症（abnormal hemoglobin）；另一类是珠蛋白基因缺失或缺陷导致珠蛋白肽链合成速率降低而产生的珠蛋白生成障碍性贫血（地中海贫血，thalassemia）。

全世界至少1.5亿人携带血红蛋白病基因，主要分布于非洲、地中海和东南亚地区。我国血红蛋白病的总发病率为0.24%～0.33%，以南方地区发病率为高。

（一）血红蛋白及其遗传控制

1.血红蛋白的组成、种类和发育演变　血红蛋白（hemoglobin，Hb）是由珠蛋白和血红素组成的结合蛋白。每一个珠蛋白分子由2条α链（或类α链，即ζ链）和2条非α链（β链或类β链，即ε链、γ链和δ链）构成，α链由141个氨基酸组成，非α链由146个氨基酸组成。每条链结合1个血红素构成1个单体，4个单体聚合成1个球状血红蛋白分子。6种不同的珠蛋白链组成人类的6种不同的血红蛋白，即Hb Gower1（$\zeta_2\varepsilon_2$）、Hb Gower2（$\alpha_2\varepsilon_2$）、Hb Portland（$\zeta_2\gamma_2$）、Hb F（$\alpha_2\gamma_2$）、Hb A（$\alpha_2\beta_2$）和Hb A_2（$\alpha_2\delta_2$）。由于γ链有2种亚型：第136位为甘氨酸的小链称$^G\gamma$，第136位为丙氨酸的小链称$^A\gamma$。因此，HbF（$\alpha_2{}^A\gamma_2$、$\alpha_2{}^G\gamma_2$）、Mb portland（$\zeta_2{}^A\gamma_2$、$\zeta_2{}^G\gamma_2$）也均有2种亚型。

在人体发育的不同阶段，上述各种血红蛋白先后出现，并且有规律地相互交替。在胚胎发育早期，合成胚胎血红蛋白：Hb Gower1、Hb Gower2和Hb Portland；胎儿期（从第8周到出生）主要的血红蛋白是Hb F；成人有3种血红蛋白：Hb A，约占97%；Hb A_2，约占2%；Hb F，约占1%。在发育过程中，各种血红蛋白消长的协调，是因为不同珠蛋白链随发育的进程有规律地发生、发展和消退。

2.血红蛋白的遗传控制　人类血红蛋白的珠蛋白肽链受控于2个基因簇，一个是α珠蛋白基因簇(α globin gene cluster)，另一个是β珠蛋白基因簇(β globin gene cluster)。

(1)α珠蛋白基因簇　人类α珠蛋白基因簇位于16号染色体短臂(16p13.33－p13.11)，每条16号染色体上有3个有功能的α珠蛋白基因(ζ、α_2、α_1)及4个假基因($\psi\zeta_1$、$\psi\zeta_2$、$\psi\alpha_1$、θ)。它们紧密连锁，排列顺序是$5'\to\zeta\to\psi\zeta_1\to\psi\zeta_2\to\psi\alpha_1\to\alpha_2\to\alpha_1\to\theta\to3'$，总长度为30 kb。$\alpha_2$和$\alpha_1$基因均有3个外显子和2个内含子，内含子1(IVS1)位于第31和32密码子之间，长95 bp；内含子2(IVS2)位于第99与100密码子之间，长125 bp(图8－1)。

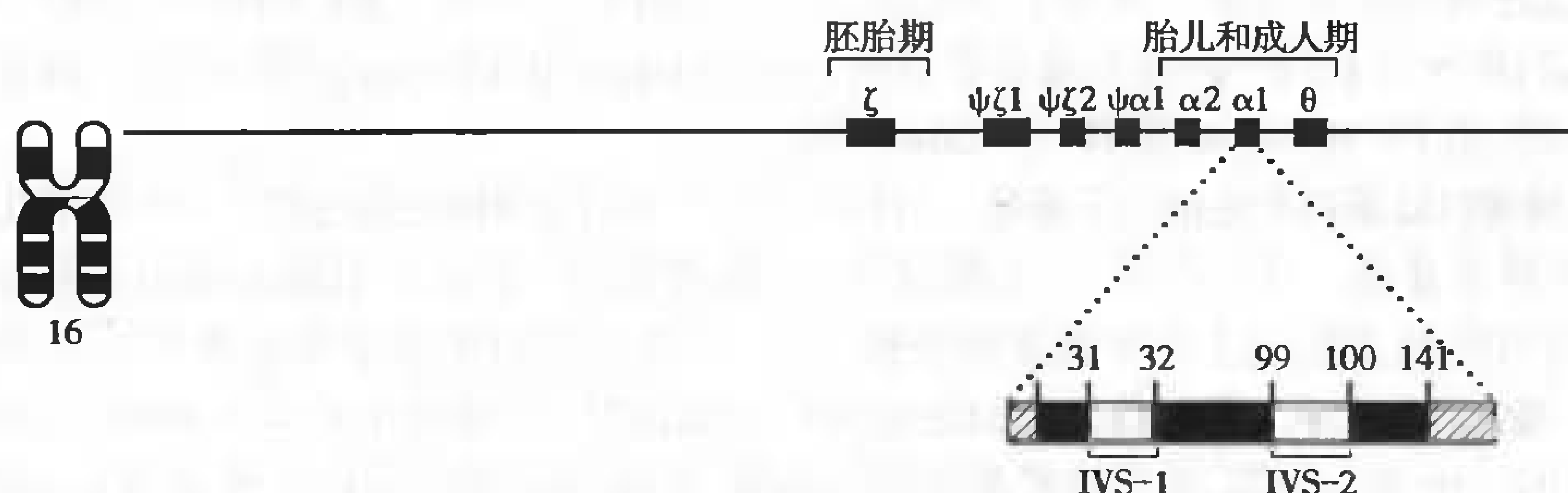


图8－1　人类α珠蛋白基因簇

(2)β珠蛋白基因簇　人类β珠蛋白基因簇位于11号染色体短臂(11p15.5)，每条11号染色体上有5个有功能的β珠蛋白基因(ε、$^G\gamma$、$^A\gamma$、δ和β)及1个假基因ψβ。它们也紧密连锁，排列顺序是$5'\to\varepsilon\to{}^G\gamma\to{}^A\gamma\to\psi\beta\to\delta\to\beta\to3'$，总长度为70 kb。β基因也有3个外显子和2个内含子，IVS1和IVS2长度分别为130 bp和850 bp，它们分别插入第30和31、104和105密码子之间(图8－2)。

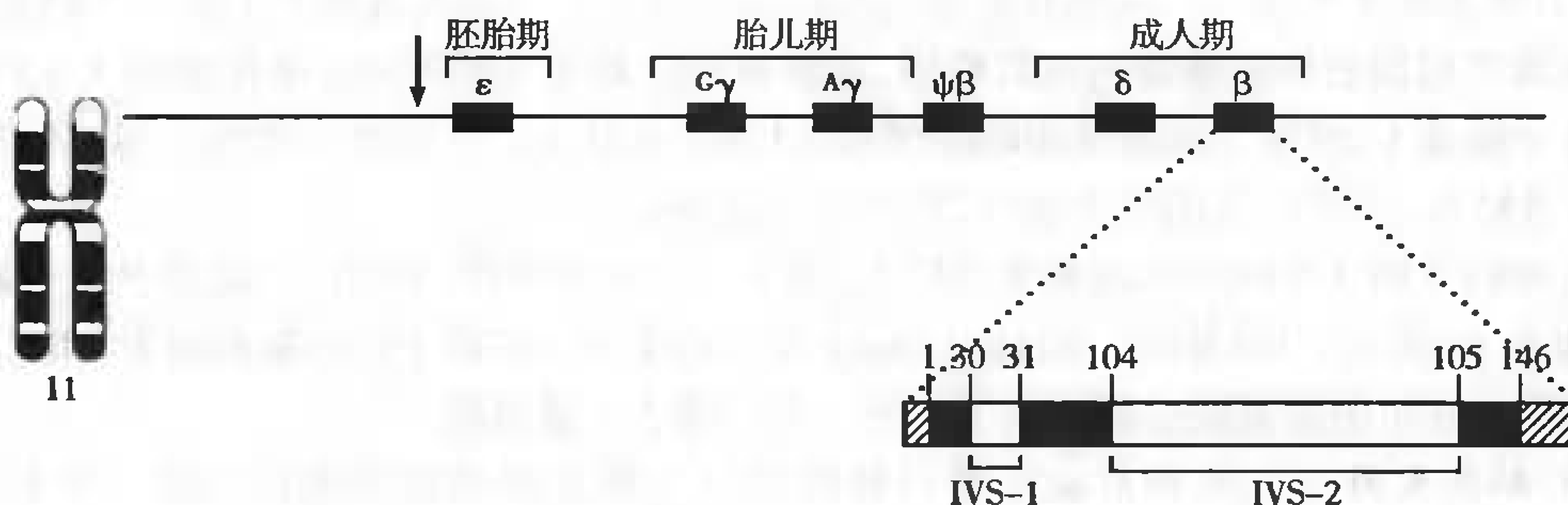


图8－2　人类β珠蛋白基因簇

DNA序列分析表明，同类的珠蛋白基因编码区具有高度的同源性。假基因没有内含子结构，75%～80%的碱基序列与正常基因相同，但由于积累了一些突变而不能翻译为结构蛋白，它们可能是进化过程中遗留下来的“退化”基因。另外，珠蛋白基因簇的排列顺序与它们在人体发育过程中的表达顺序完全一致。

(二)血红蛋白变异与异常血红蛋白症

由于编码珠蛋白肽链的基因突变,引起相应的 mRNA 改变,从而产生血红蛋白的变异。自从 1949 年 Pauling 发现镰形细胞血红蛋白 Hb S 以来,截至 2002 年,全世界已发现异常血红蛋白 681 种。

1.异常血红蛋白的命名 世界各地发现的异常血红蛋白,最初用英文字母表示,如 Hb C、Hb D、Hb E、Hb G 和 Hb M 等,一直到字母 Q。1960 年起,在改用的命名系统中,加用了发现该血红蛋白的地名命名法。原来处于电泳同一泳动位置,但后来发现是不同的变异体,这时就在字母后加上地名予以区别。因此文献上各种血红蛋白的命名,有用字母表示的,如 Hb S 和 Hb C 等;有用地名表示的,如 Hb Torino 和 Hb Sydney 等;也有字母和地名一起用的,如 Hb M Boston 和 Hb G Chinese 等。

2.异常血红蛋白产生的分子基础 引起异常血红蛋白的基因突变方式主要有以下几种。

(1)碱基替换 又称点突变,为基因突变中最普遍的一种。在目前发现的异常血红蛋白中,约 90%属于肽链上单个氨基酸替换,这是由于相应的密码子发生单个碱基替换所引起。例如,Hb S β 链第 6 位正常的谷氨酸被缬氨酸替代,是由于相应的密码子 GAG 变成了 GTG。Hb E β 链第 26 位赖氨酸取代谷氨酸,是相应的密码子 GAG 变成 AAG 所致。

单个碱基替换若发生在终止密码,使终止密码变成氨基酸的密码子,或使一个编码氨基酸的密码变成一个终止密码,则会使肽链增长或缩短,从而形成异常血红蛋白。如 α 基因第 142 位的终止密码 UAA 若突变为 CAA,即成为谷氨酰胺的密码子,则在第 142 位翻译不停止,一直进行到 173 位,产生 172 个氨基酸的 α 链,这种 α 链构成的血红蛋白称 Hb Constant Spring。又如,正常 β 基因终止密码在第 147 位,如果第 145 位的密码子 UAU(酪氨酸)突变为 UAA(终止密码),则只能合成含 144 个氨基酸的 β 链,这种 β 链和正常的 α 链构成 Hb Mckees - Rock。

(2)碱基缺失或插入 如果在珠蛋白基因的碱基顺序中丢失或插入 1 或 2 个碱基,将导致突变部位以后的碱基排列依次位移,重新编码。如在 α 基因第 138 位密码子 UCC 中缺失 1 个碱基 C,致使 3′端碱基编码顺序依次位移,而且 142 位的终止密码亦因移码而改变到了 147 位,由此产生的异常血红蛋白称为 Hb Wayne。

若密码子的 3 个碱基同时缺失或插入,则不引起读码顺序的变化,只是使肽链中减少或增加 1 个(或几个)氨基酸。如 Hb Leinden 是 β 链第 6 位或第 7 位谷氨酸缺失,Hb Grady 是 α 链第 116 位的脯氨酸后面插入了苯丙 - 苏 - 脯 3 个氨基酸。

(3)融合基因 一些异常血红蛋白肽链是由 2 种不同的肽链连接而成。例如,Hb Lepore 的非 α 链 N 端与 δ 链相同,C 端与 β 链相同,称 δβ 链;与此相反,Hb anti - Lepore 的 N 端与 β 链相同,C 端与 δ 链相同,称 βδ 链。这是由于减数分裂时同源染色体之间错位配对引起不等交换所致,结果导致一条染色体上含 δβ 基因,另一条染色体上含 βδ 基因。这种 2 个非同源基因部分片段拼接而成的基因称为融合基因(图 8 - 3)。

3.异常血红蛋白症 血红蛋白变异后,有些并不致病,仅在群体筛查时才被发现;有些血红蛋白变异体引起功能改变,如溶解度降低(如 Hb S)、对氧的亲和力增高(如 Hb Chesapeake)或降低(如 Hb Kansas)、或形成不能携带氧的高铁血红蛋白(如 Hb Boston),从而产生异常血红蛋白症。异常血红蛋白症的主要临床表现是溶血性贫血、红细胞代偿性

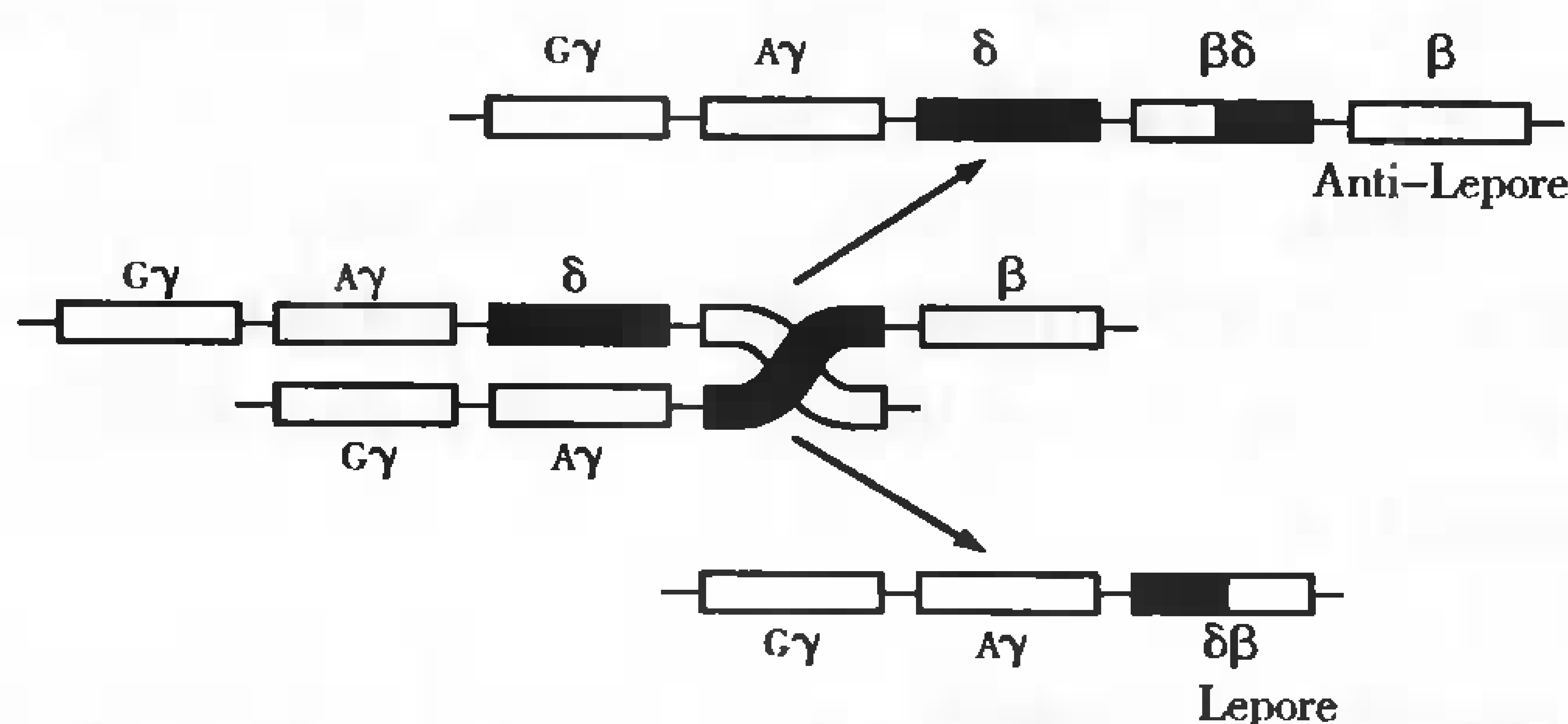


图 8－3　融合基因形成

增多和紫绀。

（1）镰形细胞贫血症　镰形细胞贫血症（sickle cell anemia）是因 β 珠蛋白基因点突变引起的一种疾病。患者血红蛋白 β 链第 6 位谷氨酸被缬氨酸取代，形成异常血红蛋白 Hb S。这种血红蛋白分子表面电荷改变，出现一个疏水区，导致溶解度降低。在氧分压低的静脉中 Hb S 凝成结晶状，使红细胞呈镰刀状（图 8－4）。镰变细胞使血液黏性增加，易使微血管栓塞，造成局部组织缺氧，甚至坏死，导致肌肉、骨关节或腹部剧痛。同时镰形细胞变形能力降低，通过狭窄毛细血管时易破裂，导致溶血性贫血。本病遗传方式为常染色体隐性遗传，杂合子（Hb^AHb^S）大部分无临床症状，但在氧分压低时可产生部分红细胞镰变；少数有轻度慢性贫血表现。以 β 珠蛋白基因相关序列为探针作分子杂交、或由限制性片段长度多态性（RFLP）分析可检出该病杂合子或进行产前诊断（图 14－1）。

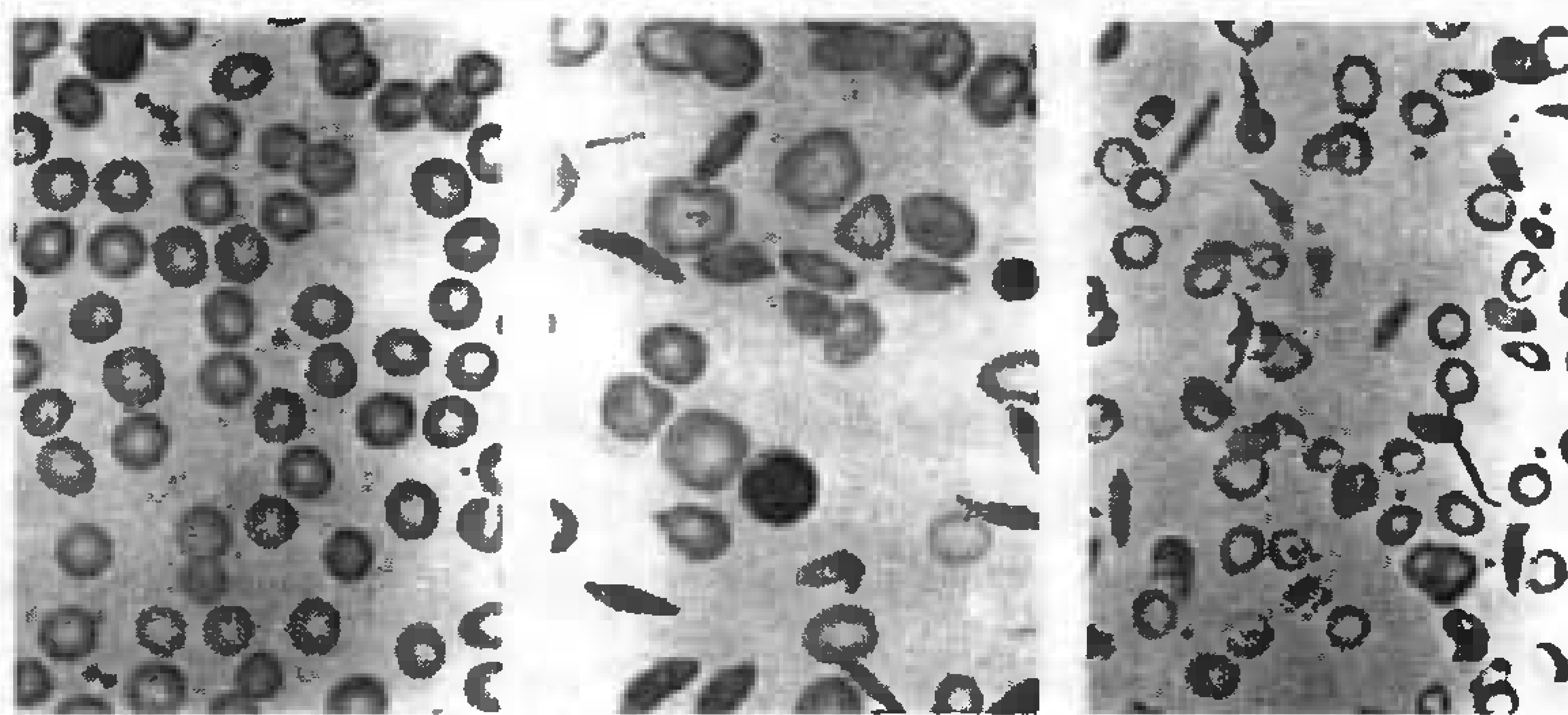


图 8－4　正常红细胞和镰形红细胞
A. 正常红细胞；B. 镰形红细胞；C. 镰形红细胞危象

（2）Hb M 病　此病又称高铁血红蛋白症。血红蛋白（Hb A）血红素中的铁原子与珠蛋白链上特定的组氨酸连接和作用，保证二价铁离子（Fe^{2+}）的稳定，以维持与氧的亲和力。血红蛋白 M 病患者珠蛋白基因点突变，使珠蛋白链与铁原子连接的组氨酸或邻近的氨基酸发生替代，导致部分血红素的 Fe^{2+} 变成 Fe^{3+}，形成高铁血红蛋白，从而丧失与氧的结合能力，使组织细胞供氧不足。患者有紫绀及继发性红细胞增多的表现。Hb M 病呈常染色体显性遗传，杂合子 Hb M 含量在 30% 即出现症状；由于纯合子患者成活率低，故很难见到本病纯合子患者。

(3)Hb Bristol 不稳定血红蛋白病　本病是由于血红蛋白β链67位缬氨酸被天冬氨酸取代，导致血红蛋白分子稳定性降低所致。这种不稳定血红蛋白易变性而沉淀于细胞中，形成变性珠蛋白小体(Heinz 小体)，含有 Heinz 小体的红细胞阳离子通透性增加、变形性降低，当通过微循环时，红细胞被阻留破坏，导致溶血。患者通常有先天性溶血性贫血、黄疸、脾肿大等临床表现。本病一般为常染色体不完全显性遗传，患者多为杂合子。重症者可发生溶血危象而危及生命。

(三)地中海贫血

地中海贫血也称为珠蛋白生成障碍性贫血，是由于珠蛋白基因缺失或缺陷，导致某种珠蛋白肽链的合成速率降低或完全不能合成，造成一些肽链缺乏，另一些肽链相对过多所致，即α链和非α链不平衡而产生的溶血性贫血。患者红细胞内多余的珠蛋白链沉积在红细胞膜上，改变膜的通透性和柔韧性，引起溶血性贫血。地中海贫血分为2种主要类型：α链合成减少或缺如的称α地中海贫血，β链合成减少或缺如的称β地中海贫血。

1.α地中海贫血　α地中海贫血主要是α基因缺失造成。α链由2对α基因控制，一条16号染色体上的2个α基因都缺失的单倍型称α^0地贫(或α地贫1)，一条16号染色体上的2个α基因之一缺失的单倍型称α^+地贫(或α地贫2)。这2种单倍型的不同组合表现为α地中海贫血的不同临床类型。

(1)Hb Bart 胎儿水肿综合征　患病胎儿2条16号染色体上的4个α基因全部缺失，为α^0地贫纯合子(α^0/α^0)，机体完全不能合成α链，以致不能形成胎儿期的正常血红蛋白HbF，过量的γ链形成 Hb Bart($^A\gamma_4$、$^G\gamma_4$)。Hb Bart 对氧的亲和力很高，在氧分压低的组织中，不易释放氧，造成组织缺氧而致胎儿水肿。这种 Hb Bart 胎儿水肿综合征胎儿多于妊娠30~40周死亡或早产，且早产儿也常在产后半小时内死亡。

(2)Hb H 病　患者为α^0地贫和α^+地贫杂合子(α^0/α^+)，即4个α基因中有3个缺失，只有少量α链合成，过量的β链聚合成 Hb H(β_4)。Hb H 不稳定，容易解聚、沉淀以致红细胞破坏，导致轻、中度贫血。

(3)标准型α地中海贫血　该型患者缺失2个α基因，基因型为α^0地贫杂合子(α^0/α^A)或α^+地贫纯合子(α^+/α^+)。由于患者能合成一定量的α链，所以仅表现轻度溶血性贫血，但如双亲均为α^0地贫杂合子(α^0/α^A)，所生子女有1/4可能性患 Hb Bart 胎儿水肿综合征。

(4)静止型α地中海贫血　该型为α^+杂合子(α^+/α^A)，即4个α基因中只有1个缺失，临床上无症状，仅在出生时血液中可检出1%~2%的 Hb Bart。此型个体与标准型α地贫者(α^0/α^A)婚配，有1/4可能性生育 Hb H 病患儿。

2.β地中海贫血　β地中海贫血的主要特征是β基因控制的β链合成缺陷，并持续合成γ和δ链。β地中海贫血主要有2类：完全不能合成β链的称β^0地贫，能部分合成β链的称β^+地贫。根据临床表现，将β地中海贫血分为重型、中间型和轻型。

(1)重型β地中海贫血　患者基因型为β^0/β^0或β^0/β^+，前者完全不能合成β链，后者β链合成量很少，造成α链大量"过剩"而沉积于红细胞膜上，引起严重溶血反应，同时γ链代偿性表达增加使 Hb F 升高(可达50%~90%)。患儿出生后几个月便可出现严重的进行性溶血性贫血、肝脾肿大等；由于骨髓增生、骨质疏松，可出现鼻塌眼肿、上颌前突、头大额隆等特殊的"地中海贫血面容"。

(2)中间型β地中海贫血　这类患者通常是某种β地贫变异型的纯合子，如β^+(高F)/β^+(高F)，或2种不同变异型地贫的双重杂合子，如$\beta^+/\delta\beta^+$。患者的临床表现介于重型和轻型之间，故称中间型地中海贫血。

(3)轻型β地中海贫血　患者基因型为β^0/β^A或β^+/β^A，这类患者可以合成相当量的β链，只有轻度贫血，甚至可由于代偿而无症状。本病特点是Hb A_2和Hb F升高。

β地中海贫血不像α地中海贫血那样是由于大片段基因缺失所致，而多数是由于基因突变引起转录翻译障碍或转录产物加工缺陷所致。

二、血浆蛋白病

血浆蛋白是血液中含量高、种类多、功能重要的一类蛋白质，在体内起着物质运输、止血、凝血免疫防御等作用。人体血浆蛋白基因异常导致血浆蛋白病。

(一)血友病

血友病(hemophilia)是一组遗传性出血性疾病，以血液中某些凝血因子的遗传性缺乏所致的严重凝血功能障碍为主要特征。根据缺乏的凝血因子的不同，血友病分为3型：甲型、乙型和丙型。丙型较少见。

1.甲型血友病　甲型血友病(hemophilia A)是凝血因子Ⅷ遗传性缺乏所致。凝血因子Ⅷ是一种复合分子，由抗血友病球蛋白(antihemophilic globulin，AHG)、因子Ⅷ相关抗原和促血小板黏附因子3种成分组成。其中AHG为因子Ⅷ凝血成分，因子Ⅷ相关抗原是因子Ⅷ凝血活性的载体蛋白，促血小板黏附因子促进血小板黏附于血管壁。甲型血友病患者因子Ⅷ凝血成分低或无，但因子Ⅷ相关抗原正常或增高。

本病的主要临床表现是反复自发性或轻微损伤后出血不止和血肿压迫引起的并发症。根据凝血因子Ⅷ的促凝活性和病情的严重程度，血友病A可分为3型：①重型，因子Ⅷ促凝活性小于20 U/L，患者出生后即发病，有自发性肌肉、关节出血，发病频繁；②中间型，因子Ⅷ促凝活性大于20 U/L，小于50 U/L，发病年龄较早，出血倾向明显；③轻型，因子Ⅷ促凝活性大于50 U/L，小于250 U/L，发病年龄较晚，无自发性出血，关节、肌肉出血情况也较少。

本病为X连锁隐性遗传。*AHG*基因定位于Xq28，长186 kb，为一巨大基因，有26个外显子。基因突变涉及碱基取代、缺失、插入和移码。

2.乙型血友病　乙型血友病(hemophilia B)是由于凝血因子Ⅸ即血浆凝血活酶成分(PTC)遗传性缺乏所致。遗传方式与甲型血友病相同，临床表现与甲型血友病相似，但出血通常较轻。

*PTC*基因定位于Xq27.1－q27.2，长34 kb，含8个外显子，编码415个氨基酸。*PTC*基因突变也有碱基取代、缺失、插入和移码等类型。应用相关限制性酶和*PTC*基因探针进行DNA分析，可对乙型血友病进行基因诊断。近年来，该病的基因治疗研究取得了突破性进展。

(二)α_1抗胰蛋白酶缺乏症

α_1抗胰蛋白酶(α_1－antitrypsin，α_1－AT)是肝脏合成的一种糖蛋白，释放入血浆后，组

成 α_1 球蛋白的绝大部分(90%)。α_1 - AT 具有抑制血清中胰蛋白酶的作用。近年来发现,先天性 α_1 - AT 缺乏和家族性肺气肿、婴儿肝硬化有密切关系。正常人群(MM 型)α_1 - AT 活性为 100%,α_1 - AT 缺乏症纯合子(ZZ 型)个体血清中 α_1 - AT 活性只有正常人的 10% ~ 15%,这些人易患肺气肿和肝硬化;杂合子(SS 型)个体血清中 α_1 - AT 比 ZZ 型者高,但比 MM 型者低,约为 60%,故这些人也有患肺气肿和肝硬化的倾向。

α_1 - *AT* 基因定位于 14q32.1,全长 12.2 kb,含 7 个外显子,编码 418 个氨基酸。α_1 - *AT* 基因突变方式主要为碱基替换。

三、受体蛋白病

受体是存在于细胞膜上、细胞质中或核内,能接受外界信息的一类特殊蛋白质。信号分子与特异性受体结合将引起细胞一系列反应,特异地改变细胞的代谢过程。控制受体蛋白合成的基因突变,受体蛋白的质和量将改变,继而影响细胞代谢而致病。受体蛋白遗传性缺陷引起的疾病称为受体蛋白病(receptor protein disease)。

(一)家族性高胆固醇血症

家族性高胆固醇血症(familiar hypercholesterolemia, FH)为高脂蛋白血症Ⅱ型,是基因突变导致细胞膜上低密度脂蛋白受体(low density lipoprotein receptor, LDLR)缺陷所引起。在正常情况下,LDL 与 LDLR 结合后,经内吞作用进入细胞,被溶酶体酶水解,释放出游离胆固醇。游离胆固醇抑制微粒体上的 β—羟基—β—甲基戊二酰辅酶 A 还原酶(HMG - CoA 还原酶)活性,使内源性胆固醇合成降低;同时游离胆固醇激活胆固醇酯化酶,促使胆固醇酯化为胆固醇酯贮存起来(图 8 - 5)。当 LDLR 缺陷时,LDL 不能进入细胞,使细胞内胆固醇合成反馈调节受阻,胆固醇合成不受抑制,胆固醇酯生成减少,从而导致细胞内和血液中游离胆固醇过多,导致高胆固醇血症。

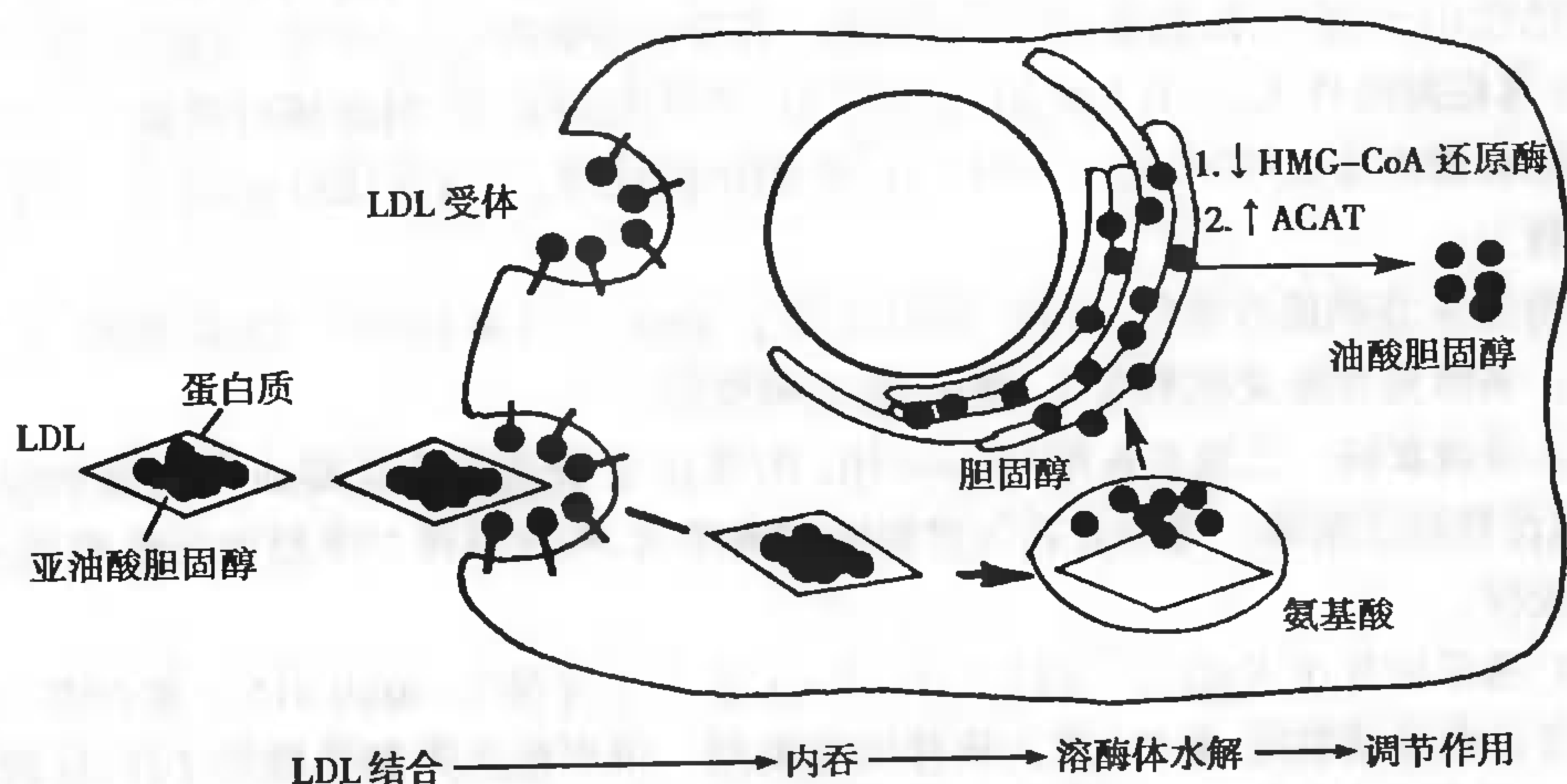


图 8 - 5　低密度脂蛋白受体作用示意图

本病为常染色体不完全显性遗传。人群中杂合子患者约为 1/500,其血清胆固醇为 3 ~ 4 g/L,手、肘、膝、踝部可有黄瘤,并有角膜弓,40 ~ 60 岁可发生冠心病;纯合子罕见,其

血清胆固醇超过 6 g/L,幼年即可出现黄瘤和角膜弓,20 岁前发生冠心病。

LDLR 基因定位于 19p13.3 - p13.1,长约 45 kb,含 18 个外显子。目前已发现的 *LDLR* 基因突变 20 多种,缺失型突变导致 LDLR 完全不能合成,其他类型的突变导致 LDLR 不能转运到细胞膜上,或 LDLR 不能与 LDL 结合,或 LDL 与 LDLR 结合后不能被内吞等。

(二)睾丸女性化综合征

睾丸女性化综合征(testicular feminigation syndrome)是一种 X 连锁隐性遗传病。患者外观为女性,青春期后乳房、外阴均发育良好。常因原发性闭经、不育等原因求治时被发现。体格检查可见阴蒂肥大、阴道短浅、子宫、卵巢缺如,腹腔或腹股沟处或大阴唇皮下有睾丸和附睾,睾丸曲细精管不能生成精子。染色体核型分析结果为 46,XY。约 8%患者青春期睾丸恶变,且随年龄增大,恶变概率增高。

本病一般认为是由于基因突变导致细胞质内雄性激素受体缺乏所致,患者虽有睾丸,亦能正常产生雄激素,但雄激素不能在靶器官发生作用,所以表现女性化特征。雄激素受体基因(*TFM*)位于 Xq13 和 Xp11 之间。

四、膜转运载体蛋白病

一些小分子物质进出细胞,需通过细胞膜的特异性主动转运系统。在这种转运系统中,如果发生载体蛋白的缺陷,必然影响物质代谢,从而引起疾病。由于基因突变导致的膜载体蛋白缺陷引起的疾病称为膜转运载体蛋白病(transport protein deficiency)。

(一)肝豆状核变性

肝豆状核变性(hepatolenticular degeneration)又称 Wilson 病(Wilson disease, WD),是一种铜代谢障碍导致的疾病。该病主要特点是肝硬化伴基底神经节豆状核变性,呈常染色体隐性遗传。群体发病率为 1/16 万,杂合子频率为 1/200。

本病患者细胞膜与铜转运有关的 ATP7B 缺陷,使铜不能被及时清除,沉淀于组织细胞而引起毒性作用。受累器官以肝、脑、肾等最为显著。肝中铜的浓度升高,引起肝坏死,最终发展为肝硬化;铜沉淀于肾,使近曲小管受损而出现氨基酸尿、蛋白尿等;沉积于脑导致神经系统受损;沉积于角膜形成角膜外缘绿色环;铜被红细胞摄取,可发生溶血性贫血。

本病发病年龄多在 10~25 岁,最初表现类似慢性活动型肝炎,继而发展为肝硬化,出现黄疸,肝脾肿大,腹腔积液等严重肝功能衰竭的表现。约 40%患者以神经系统临床表现为主,表现为发音和吞咽困难、运动失调、步态不稳、扑翼样震颤,偶有癫痫发作。头颅 CT 检查可见脑萎缩、豆状核低密度灶、脑室扩大等。

本病致病基因 *WD* 定位于 13q14.3。

(二)胱氨酸尿症

胱氨酸尿症(cystinuria)是肾小管重吸收胱氨酸、赖氨酸、精氨酸和鸟氨酸的特异性载体蛋白缺陷所产生的疾病。遗传方式为 AR。患者血浆中这 4 种氨基酸低于正常,尿液中水平增高。患者小肠黏膜上皮细胞的主动转运可能也有类似缺陷。由于这 4 种氨基酸中赖氨酸、精氨酸和鸟氨酸易溶于水,而胱氨酸不易溶于水,所以当患者胱氨酸每日排出量达 0.5~1.0 g 或超过饱和浓度(pH5~7 时尿中胱氨酸饱和浓度为 0.3~0.4 g/L)时,则尿液中形成胱氨酸结晶,日久形成尿道结石。本病患者主要临床表现是尿道结石及由此引

起的尿路感染和肾绞痛。

研究表明,胱氨酸尿症至少涉及3个异常等位基因,因而,该病呈现3种不同临床类型。

Ⅰ型:纯合子患者4种氨基酸排出量均增加,杂合子氨基酸排出量正常,患者肠上皮细胞对4种氨基酸均不能吸收。该型相关基因定位在2p16.2。

Ⅱ型:纯合子患者4种氨基酸排出量均增加,杂合子的胱氨酸和赖氨酸排出少量增加。纯合子患者肠上皮细胞对4种氨基酸有少量摄取。该型相关基因定位在2p16.3。

Ⅲ型与Ⅱ型相似,纯合子患者肠上皮细胞对氨基酸吸收稍低于正常。该型相关基因定位在19q13.1。

第二节 先天性代谢病

先天性代谢病是酶的遗传性缺陷所引起。酶的种类很多,因此先天性代谢病的种类亦很多,至今已发现数千种,大多数为常染色体隐性遗传病,少数为X连锁隐性遗传病。

一、先天性代谢病发生的基本原理

酶是生物催化剂,基因通过合成特定的酶来控制机体的新陈代谢,进而影响遗传性状的形成。代谢底物进入细胞后,在一系列酶的催化下,经过一系列中间产物,最后形成代谢终产物。假定人体内的A物质在E_{AB}、E_{BC}和E_{CD} 3种酶的催化下,经B、C阶段,最后形成产物D(图8-6)。

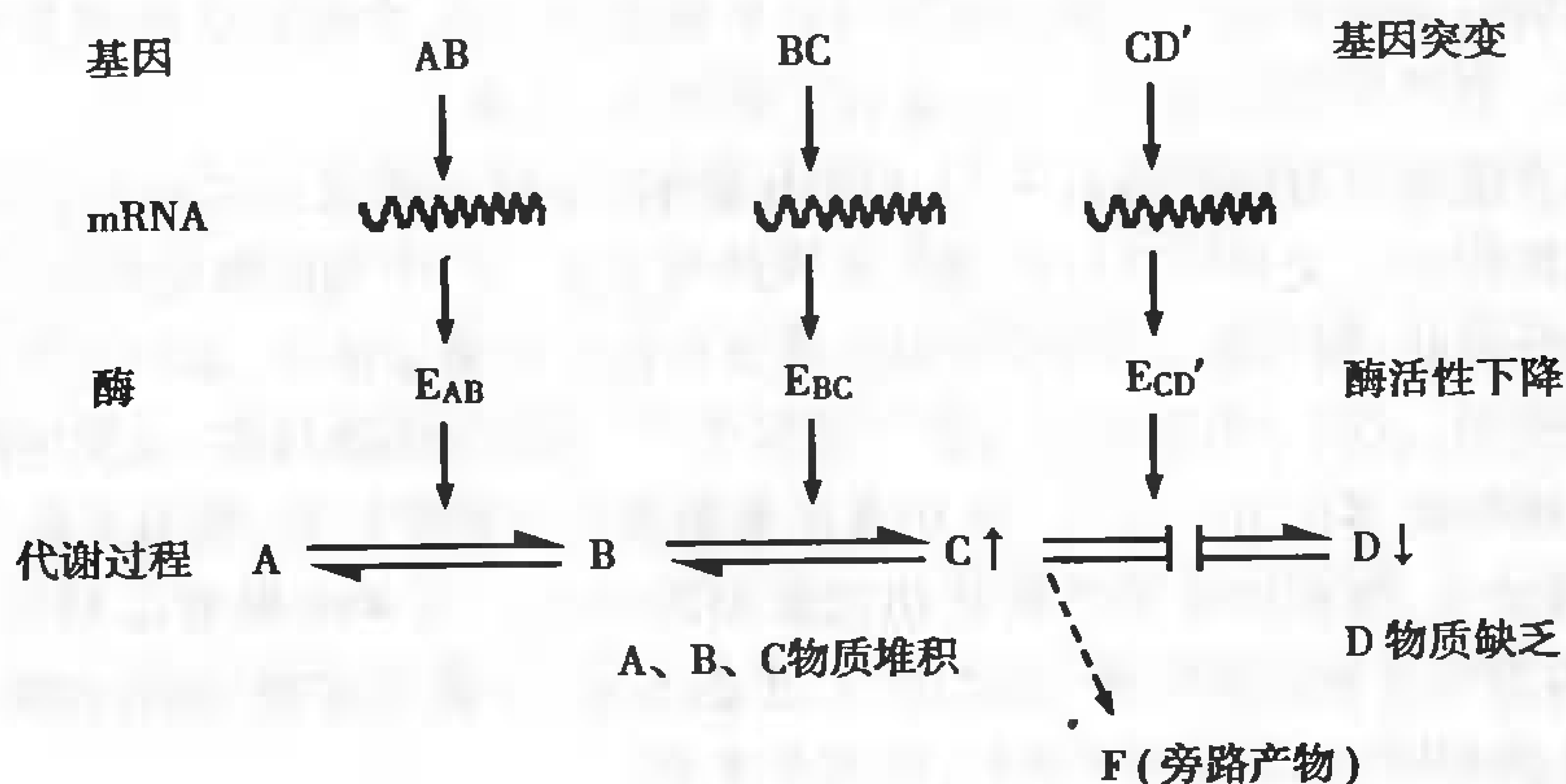


图8-6 基因、mRNA、酶和代谢过程的关系

如果基因*CD*突变为*CD′*,则其所控制的酶异常($E_{CD'}$),C→D反应不能顺利进行或完全停止,导致代谢的正常产物(D)缺乏,代谢中间产物(B、C)增多,代谢底物(A)堆积及代谢途径转向等。此外,还可使代谢的反馈调节紊乱,从而引起相应的临床表现。这就是基因突变导致先天性代谢病发病的基本机制。

二、氨基酸代谢病

氨基酸代谢病是氨基酸分解代谢过程中遗传性酶缺乏所引起的氨基酸代谢缺陷。

(一)苯丙酮尿症

苯丙酮尿症(phenylketonuria,PKU)是一种比较常见的氨基酸代谢病。据国内 11 省市新生儿筛查结果,发病率为 1/16 500。本病分为经典型和非经典型 2 种类型。

1.经典型苯丙酮尿症 是由于患者肝中缺乏苯丙氨酸羟化酶(phenylalanine hydroxylase,PAH)所引起。苯丙氨酸和酪氨酸的代谢如图 8-7 所示。苯丙氨酸经苯丙氨酸羟化酶催化可形成酪氨酸。此酶缺乏使苯丙氨酸不能转变为酪氨酸,致使苯丙氨酸在体内积累。过量的苯丙氨酸使旁路代谢活跃,产生大量苯丙酮酸及其衍生物苯乳酸、苯乙酸等。这些旁路产物由尿液和汗液排出,使患儿的毛发、皮肤和尿液均有特殊气味。过量的苯丙氨酸抑制酪氨酸脱羧酶活性,影响去甲肾上腺素和肾上腺素合成,也减少黑色素的合成。旁路产物抑制 *L*-谷氨酸脱羟酶活性,影响 γ-氨基丁酸的生成;同时还抑制 5-羟色胺脱羧酶活性,使 5-羟色胺生成减少,从而导致脑发育障碍。

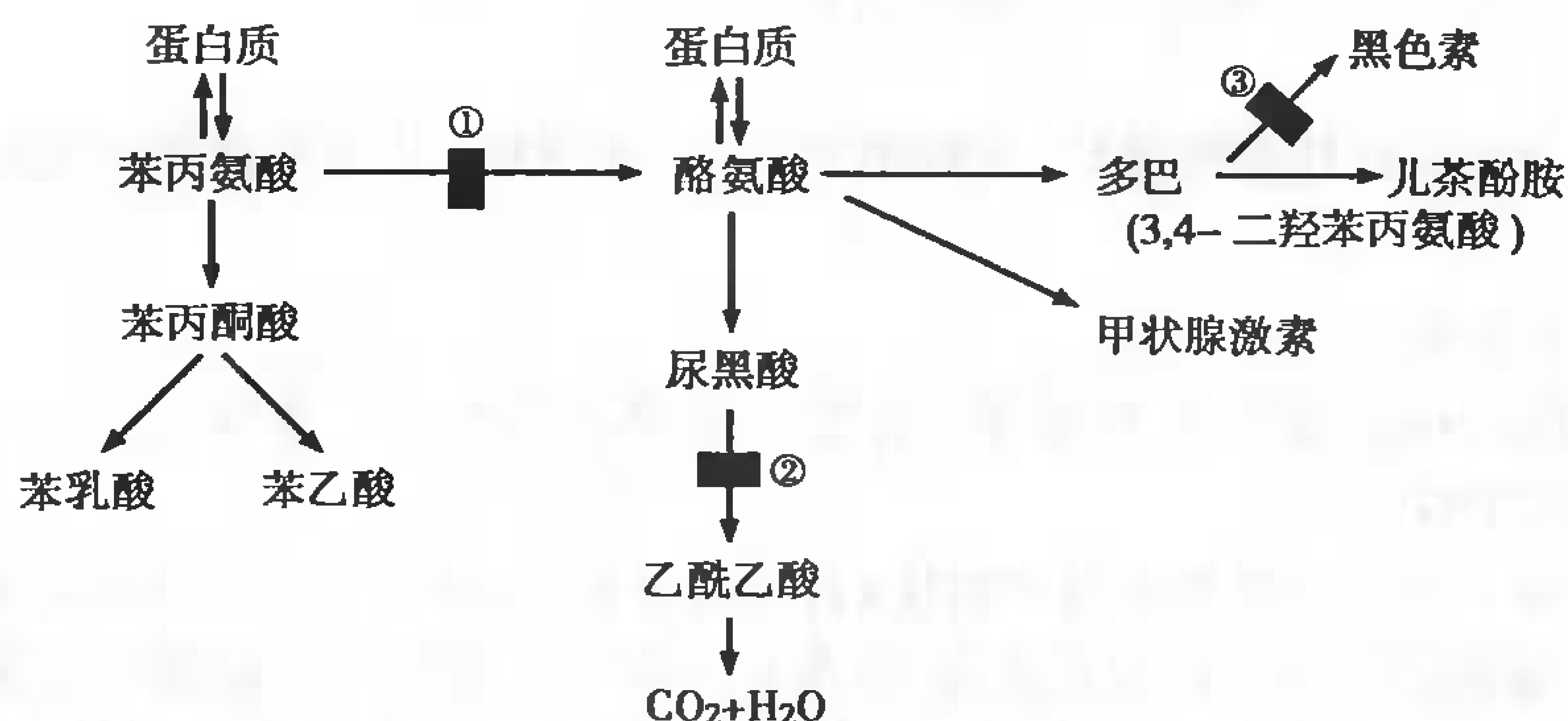


图 8-7 苯丙氨酸及酪氨酸代谢图解

①苯丙氨酸羟化酶缺乏——苯丙酮尿症;②尿黑酸氧化酶缺乏——尿黑酸尿症;③酪氨酸酶缺乏——白化病

患儿出生时无显著异常,3~4 个月后开始出现智力障碍,未治疗者将发展为痴呆。90%以上的患儿毛发淡黄、皮肤白、虹膜可呈黄色。半岁以上时出现易激怒、肌张力亢进、共济失调、震颤等,多数有脑电图异常。尿有特殊的"发霉"气味。如能早期明确诊断,给予低苯丙氨酸饮食等治疗,可控制该病病情的发展。

经典型苯丙酮尿症为常染色体隐性遗传。苯丙氨酸羟化酶基因定位于 12q24.1,长度约为 90 kb,有 13 个外显子。绝大部分苯丙氨酸羟化酶分子的缺陷来自于基因内单个碱基的替换,少数为外显子缺失。

2.非经典型苯丙酮尿症 是由二氢喋呤还原酶缺乏所致。PAH 催化苯丙氨酸生成酪氨酸的反应需要辅助因子四氢生物喋呤(BH_4)作为电子供体,在化学反应中,BH_4 被转变为二氢生物喋呤(BH_2),然后在 BH_2 还原酶催化下还原为 BH_4(图 8-8)。因此,BH_2 还原

酶缺乏将引起 BH_4 缺乏，在这种情况下，虽有正常PAH，也不能将苯丙氨酸羟化为酪氨酸，致使苯丙氨酸在体内积累，引起苯丙酮尿症。非经典型苯丙酮尿症的临床表现与经典型苯丙酮尿症相似。BH_2 还原酶基因定位于4p15.1－p16.1。

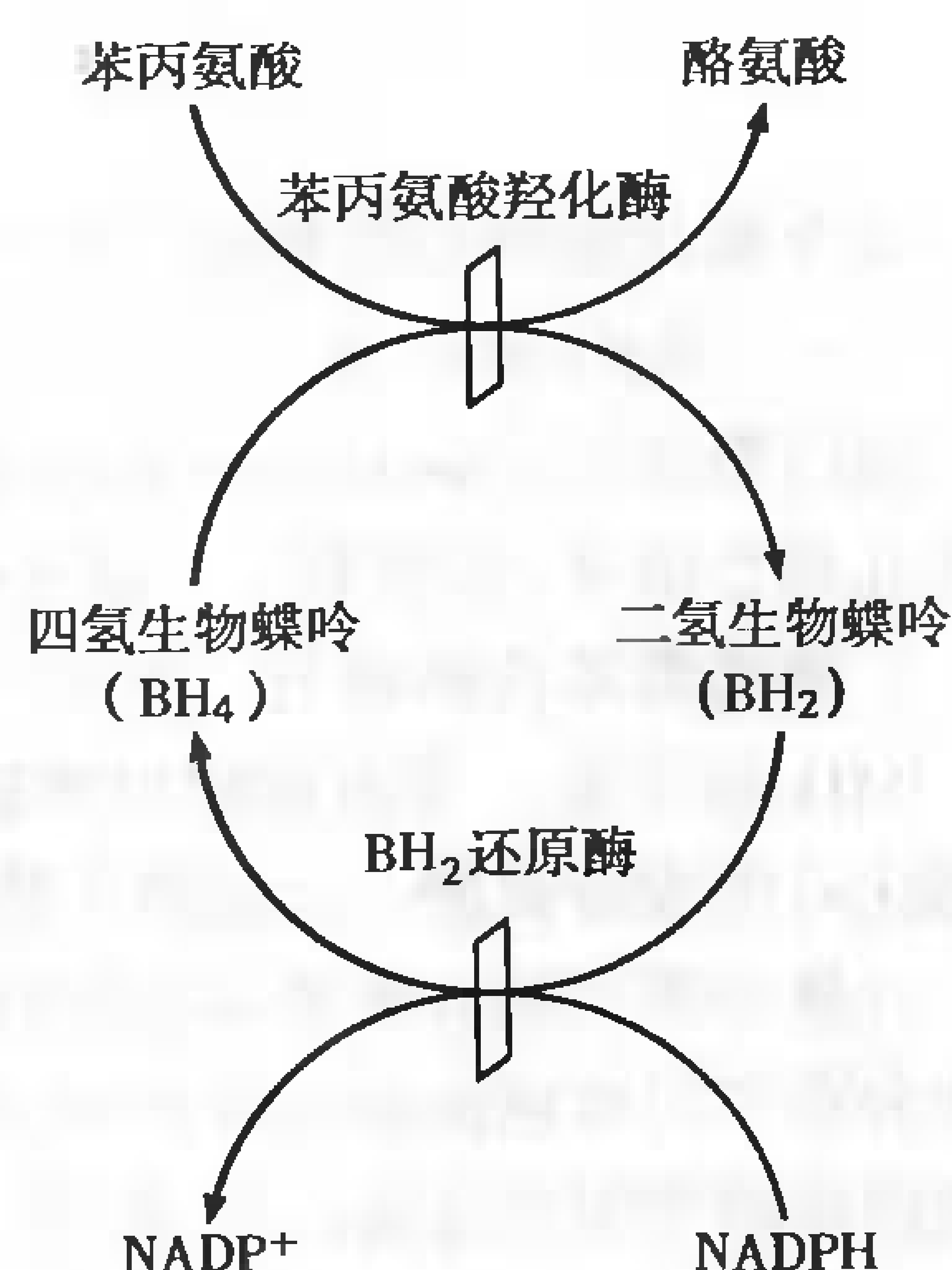


图8－8　苯丙氨酸羟化系统

(二)尿黑酸尿症

尿黑酸尿症(alcaptonuria)是尿黑酸氧化酶(homogentisic acid oxidase)先天性缺乏所致。由于尿黑酸氧化酶缺乏，尿黑酸不能氧化成乙酰乙酸和延胡索酸，结果大量尿黑酸从尿中排出。新生儿期患儿尿布上可发现紫褐色斑点，日久尿布渐呈黑褐色；儿童期除尿中排出尿黑酸外，并无其他异常；成人期除尿黑酸尿外，体内增多的尿黑酸沉淀于结缔组织，导致褐黄病，表现为皮肤、耳廓、面颊、巩膜等处弥漫性色素沉着。如果累及关节则形成褐黄病性关节炎。

本病也为常染色体隐性遗传，发病率约为1/250 000。已知尿黑酸氧化酶基因定位于3q21－q23。

(三)白化病

白化病(albinism)是先天性皮肤、毛发色素缺乏的疾病。本病分布广泛，发病率为1/10 000～1/20 000。

白化病分2型。Ⅰ型为全身性白化病，为常见类型，遗传方式为AR，相关基因定位在11q14－q21，基因全长50 kb，现已发现20多种点突变。患者全身的黑色素细胞缺乏酪氨酸酶(tyrosinase)，不能将酪氨酸转变为黑色素，此型又称为酪氨酸酶阴性型，临床表现为皮肤白皙，毛发银白或呈淡黄色，虹膜及瞳孔浅红色，羞明，视物模糊，可有眼球震颤。日晒皮肤易灼伤，暴露的皮肤易患皮肤癌。

Ⅱ型为局部性白化病，致病基因位于15号染色体短臂上，遗传方式为AD，临床表现与Ⅰ型难于区分，色素痣是Ⅱ型区别于Ⅰ型的重要鉴别点。Ⅱ型患者酪氨酸酶阳性，故又称为酪氨酸酶阳性型。Ⅱ型的发病机制尚不明了。

三、糖代谢病

糖代谢病是由于糖类代谢(或分解)过程中遗传性酶缺乏所引起。

(一)半乳糖血症

半乳糖经肠道吸收后，在肝内经一系列酶促反应转变成葡萄糖而被组织利用。半乳糖代谢过程涉及半乳糖激酶、半乳糖—1—磷酸尿苷转移酶和半乳糖尿苷二磷酸—4—表异构酶(图8－9)，这3种酶的遗传性缺乏，导致3种不同类型的半乳糖血症(galactosemia)。它们均呈常染色体隐性遗传。

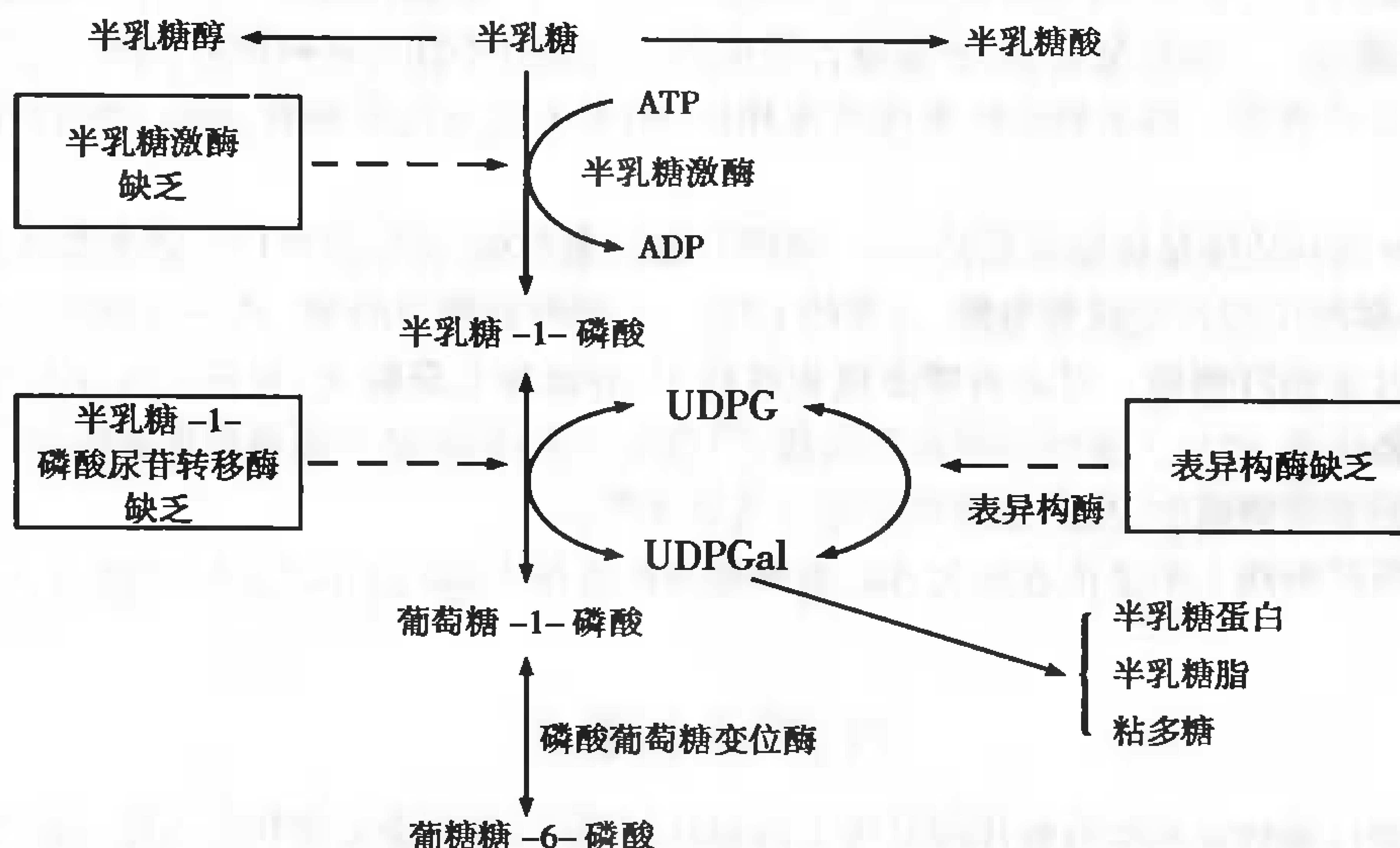


图 8-9　半乳糖代谢途径

半乳糖血症Ⅰ型又称为经典型半乳糖血症，系半乳糖—1—磷酸尿苷转移酶缺乏所致。由于半乳糖—1—磷酸尿苷转移酶缺乏，致使半乳糖—1—磷酸在肝积聚，引起肝肿大、肝功能受损；在脑中积聚可引起运动及智力障碍；在肾及肠组织积聚可导致氨基酸吸收障碍；血中半乳糖升高可使葡萄糖释出减少，出现低血糖症。另外，晶状体内半乳糖增多，激活醛糖还原酶，生成半乳糖醇；半乳糖醇能改变晶体内的渗透压，水分进入晶体内导致晶体变性，继而产生白内障。患儿出生后数日即出现呕吐、腹泻、拒食，随之发生脱水、体重下降、黄疸、肝损害，1～2 月后出现白内障。如不戒奶，几个月后出现智能障碍，病情进行性加剧。

半乳糖—1—磷酸尿苷转移酶基因定位于 9q13，现已发现多种变异型。半乳糖血症Ⅰ型遗传方式为 AR，杂合子酶活性仅为正常纯合子的一半，但不出现症状。

半乳糖血症Ⅱ型为半乳糖激酶缺乏所致。病情较半乳糖血症Ⅰ型轻，白内障常见，有时可能是惟一的临床表现，智能发育正常或迟缓，血半乳糖浓度增高，尿中出现半乳糖和半乳糖醇，但无氨基酸和蛋白。半乳糖激酶基因定位在 17q21－q22。

半乳糖血症Ⅲ型为半乳糖尿苷二磷酸—4—表异构酶缺乏所致。临床表现变化不一，可无症状或类似于Ⅰ型半乳糖血症。半乳糖尿苷二磷酸—4—表异构酶基因定位于 1p36－p35。

(二)糖原贮积病

糖原贮积病(glycogen storage disease)是一类较罕见的先天性代谢病，是与糖原分解代谢有关的酶遗传性缺陷引起糖原累积所致。不同的酶缺乏产生不同类型的糖原贮积病，现已发现 13 种类型，以Ⅰ型最常见。

糖原贮积病Ⅰ型，又称 von Gierke 病或肝肾型糖原贮积症，该病由 von Gierke 在 1929

年首次报道。患者肝、肾和肠组织完全缺乏葡萄糖—6—磷酸(glucose - 6 - phosphate,G6P)酶,G6P 酶是一个多酶复合体,根据复合体中缺陷的成分不同可将糖原贮积病Ⅰ型分为Ⅰa 和Ⅰb 2 个亚型。两亚型临床表现基本相同,病变主要累及肝和肾,但不侵犯骨骼肌和心脏。

G6P 酶的功能是催化葡萄糖—6—磷酸分解为葡萄糖,供组织利用。该酶缺乏使葡萄糖—6—磷酸不能转变成葡萄糖,过多的 G6P 一方面抑制糖原分解,另一方面通过可逆反应合成过多的肝糖原。过多的糖原沉积在肝中,导致患儿肝肿大,发育迟滞;G6P 还可循酵解途径代谢,产生大量的丙酮酸和乳酸,导致酸中毒;脂肪动员加强可出现酮血症;组织可利用的葡萄糖减少,当不进食时极易出现低血糖。

糖原贮积病Ⅰ型遗传方式为 AR,致病基因定位在 17q21,全长 12.5 kb,包含 5 个外显子。

四、脂类代谢病

脂类代谢病是脂类分解代谢过程中特异性酶缺乏,导致脂类底物在内脏、脑和血管等处累积产生的疾病,总称脂类累积症(lipidosis)。脂类结构复杂,种类多样,特异性酶缺乏所致脂类累积症也多种多样,以下简介神经鞘脂累积症的代表性疾病。

(一)Tay - Sachs 病

Tay - Sachs 病亦称 GM_2 神经节苷脂累积症(GM_2 gangliosidosis)。在正常代谢中,氨基己糖苷酶 A 分解 GM_2 神经节苷脂成为 GM_3 和 N - 乙酰氨基半乳糖。当该酶缺乏时,这一过程受阻,导致 GM_2 神经节苷脂在脑组织中蓄积引起一系列临床表现。早期表现为视网膜黄斑变性,进行性失明;局部或全身性抽搐;肌张力进行性减退,生长发育阻滞;智能低下;后期身体完全瘫痪,出现恶病质。患者平均生存期为 25.9 个月。

本病呈常染色体隐性遗传。氨基己糖苷酶 A 基因定位在 15q23 - q24,已检出的该基因突变类型有碱基取代、缺失和移码突变等。

(二)Gaucher 病

Gaucher 病(Gaucher disease)是由于葡萄糖脑苷脂酶(glucocere - brosidase)缺乏,导致葡萄糖脑苷脂蓄积于单核巨噬细胞系统的细胞中而致病,本病的特征是肝、脾、淋巴结及骨髓等处可见蓄积有葡萄糖脑苷脂的 Gaucher 细胞。

本病分为急性型和慢性型。急性型 1 岁以内起病,表现为肝、脾肿大,贫血,发育迟缓;神经系统临床表现为:意识障碍,角弓反张,四肢强直,集合性斜视吞咽困难,惊厥等。病情进展迅速,多于 1 岁前死亡。慢性型多于学龄前发病。可有肝脾肿大,贫血,骨骼受损,但神经系统多不累及。慢性型病情进展缓慢,患者可生存十几年至数十年。

葡萄糖脑苷脂酶基因定位于 1q21。采用等位基因特异性寡核苷酸探针检测 370 位 G→C 取代,并以 RFLP 检测 444 位 T→C 取代,可对大约 80% 患者进行基因诊断。通过羊水细胞测定酶活性或检测分子变异可进行该病的产前诊断。

五、嘌呤代谢病

人体内嘌呤代谢过程中酶的遗传性缺乏产生的疾病称为嘌呤代谢病,Lesch - Nyhan

综合征(Lesch－Nyhan syndrome)就属于这类疾病。该综合征是次黄嘌呤鸟嘌呤磷酸核糖转移酶(HGPRT)遗传性缺乏所致。HGPRT的功能是催化5—磷酸核糖—1—焦磷酸(PRPP)的磷酸核糖基转移至次黄嘌呤和鸟嘌呤,分别生成次黄苷酸(IMP)和鸟苷酸(GMP)。这2种核苷酸与腺苷酸可反馈抑制嘌呤的前体5—磷酸核糖—1—胺的合成。HGPRT缺陷时,这一反馈抑制作用减弱或消失,嘌呤合成加快,尿酸生成增多(图8－10)。患者红细胞和白细胞中HGPRT的含量仅为正常人的2%～10%。

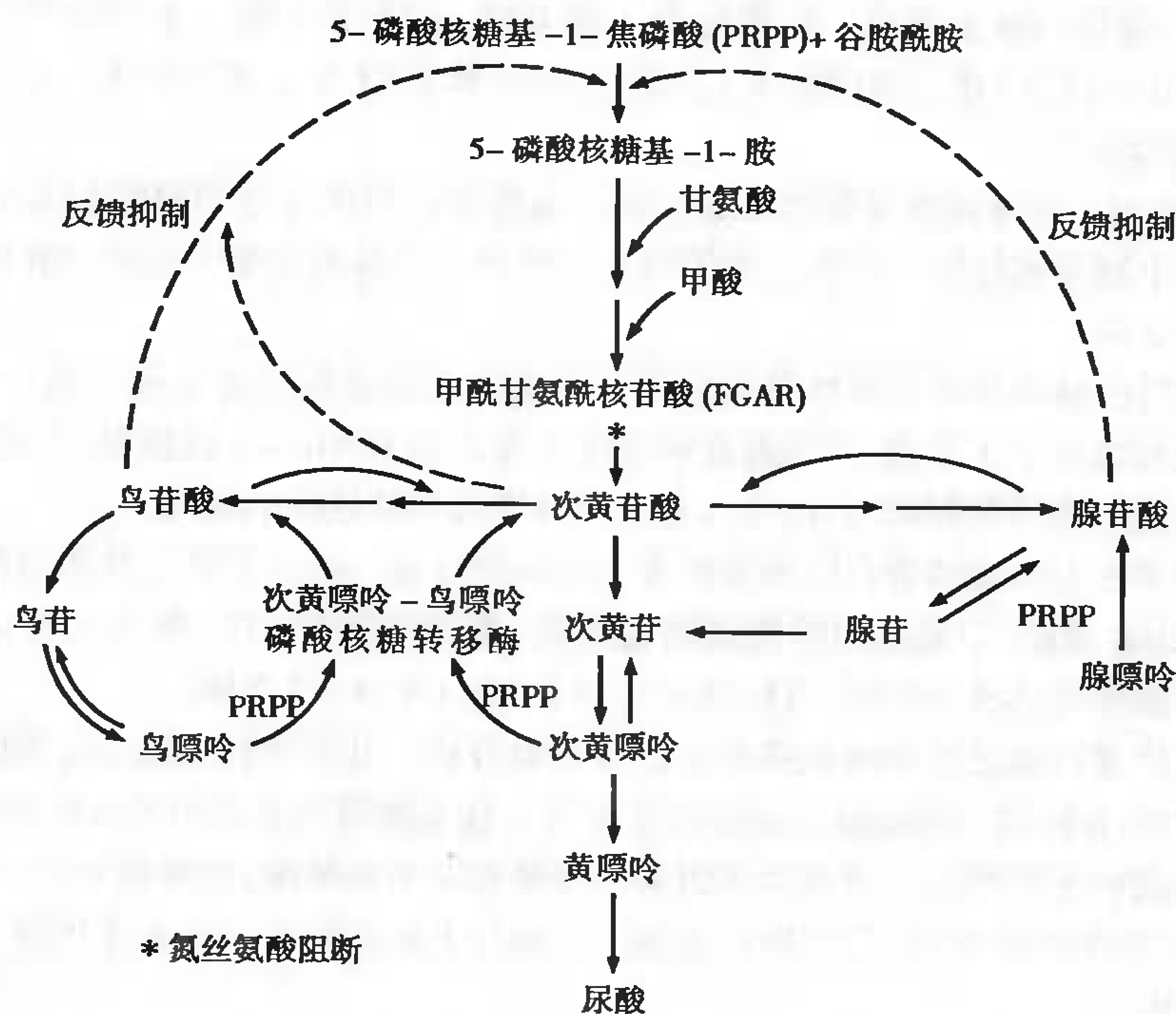


图8－10　嘌呤核苷酸合成的反馈调节

本病的临床特征是高尿酸血症和尿酸尿、尿道结石和痛风等,常伴有智能发育不全,舞蹈样动作,强迫性自残行为,因此,本病也称自毁容貌综合征。本病发病率约为1/38 000。遗传方式为XR,*HGPRT*基因定位在Xq26－q27。目前已检出本病20余种突变型,常见的突变方式有碱基取代、缺失和插入等。

小　结

分子病是基因突变导致蛋白质分子结构和功能异常而引起的疾病。由于遗传性酶缺陷导致代谢紊乱引起的疾病称为先天性代谢病。

人类的珠蛋白基因分别位于α基因簇和β基因簇,α基因簇位于16号染色体上,β基因簇位于11号染色体上。

异常血红蛋白病是珠蛋白基因异常,导致珠蛋白的肽链结构和功能异常而引起的疾病。

引起异常血红蛋白的基因突变方式有:①单个碱基替换;②碱基缺失和插入;③融合

基因形成。

典型的异常血红蛋白病包括:①镰形细胞贫血,这是β珠蛋白基因点突变导致红细胞镰变引起的贫血症;②Hb M病是由于碱基置换,使珠蛋白链氨基酸替代,形成高铁血红蛋白,从而丧失与氧的结合能力所致;③Hb Bristol不稳定血红蛋白病,是由于β链氨基酸取代,使血红蛋白稳定性降低,变性的血红蛋白沉积于细胞中引起溶血性贫血所致。

地中海贫血是α链或非α链合成速率降低或缺如,造成α链和非α链不平衡而引起。

α地贫是α基因的缺失造成。临床分为4种类型:①Hb Bart胎儿水肿综合征(4个α基因缺失);②Hb H病(3个α基因缺失);③标准型α地贫(2个α基因缺失);④静止型α地贫(1个α基因缺失)。

β地贫主要由于β基因突变导致β链合成缺陷所致。根据β链的缺陷程度和临床表现,分为重型、中间型和轻型。引起β地贫的分子机制多为基因突变引起转录翻译障碍或转录产物加工缺陷。

人体血浆蛋白基因异常导致血浆蛋白病。甲型血友病系缺乏凝血因子Ⅷ所致,乙型血友病系缺乏凝血因子Ⅸ所致,两型血友病的遗传方式均为XR。α_1抗胰蛋白酶缺乏症患者丧失了对胰蛋白酶的抑制作用,这些人有易患肺气肿和肝硬化的倾向。

受体基因突变引起受体蛋白质和量的改变而致受体蛋白病。家族性高胆固醇血症是由于细胞膜LDLR缺陷,引起胆固醇代谢紊乱所致,遗传方式为AD。睾丸女性化综合征是细胞质内雄激素受体缺乏所致,致病基因定位在Xq11和Xq13之间。

膜转运载体蛋白缺乏影响细胞膜的主动转运而致病。肝豆状核变性患者细胞膜与铜转运有关的ATP7B缺陷。胱氨酸尿症患者肾小管对胱氨酸等氨基酸的转运特异性缺陷。

先天性代谢病发病的基本原理为基因缺陷导致相应的酶缺陷,使酶催化的代谢中断,引起代谢底物、中间产物堆积,终产物缺乏而出现相应的临床异常。先天性代谢病的遗传方式大多为AR。

常见的氨基酸代谢病包括:①苯丙酮尿症,是由于苯丙氨酸羟化酶缺乏,血和尿中苯丙酮酸含量升高所致;②尿黑酸尿症,患者缺乏尿黑酸氧化酶,使尿黑酸在体内堆积并大量从汗和尿中排出;③白化病,全身性白化病患者缺乏酪氨酸酶,不能合成黑色素。

常见的糖代谢异常疾病包括:①半乳糖血症,Ⅰ型系半乳糖尿苷转移酶缺乏,基因定位于9q13;Ⅱ型为半乳糖激酶缺乏,基因定位在17q21-q22;Ⅲ型为半乳糖尿苷二磷酸—4—表异构酶缺乏,基因定位于1p36-1p35。②糖原贮积病,Ⅰ型缺乏G6P酶,基因定位于17q21。

脂类代谢病的代表性疾病有:①Tay-sachs病,是由于氨基己糖苷酶A缺乏,导致GM_2神经节苷脂在脑组织中蓄积所致。相关基因定位在15q23-24;②Gaucher病,是由于葡萄糖脑苷脂酶缺乏使葡萄糖脑苷脂蓄积所致。相关基因定位在1q21。

嘌呤代谢病的典型病例——Lesch-Nyhan综合征是HGPRT缺乏,嘌呤代谢紊乱所致,该病致病基因定位在Xq26-q27。

(税青林)

第九章 群体遗传学

群体(population),又称为种群或孟德尔式群体(Mendelian population),是指生活在某一地区并能够互相交配的同一物种的个体群。

群体遗传学是研究生物群体的遗传结构及其变化规律的遗传学分支学科,它应用数学和统计学方法来研究群体中基因频率和基因型频率以及两者之间的关系,研究突变、选择、迁移、随机遗传漂变对群体遗传结构的影响,研究群体的遗传组成及其变化规律。对遗传病的预防、监测和治疗具有积极意义。

第一节 群体中的遗传平衡

一、基因频率和基因型频率

基因频率(gene frequency)是指群体中某一基因占该基因座位上全部等位基因的比率。如果群体中某个基因座位上有 2 个基因:A 和 a,通常显性基因 A 的频率以 p 表示、隐性基因 a 的频率以 q 表示,那么,p 就是基因 A 在基因 A 和 a 总量中所占的比率,q 就是基因 a 在基因 A 和 a 总量中所占的比率,$p+q=1$。

基因型频率(genotypic frequency)是指某一基因型的所有个体在所研究群体中所占的比率。基因型频率在群体中的变化规律与基因频率不同,研究群体的遗传组成,不但要考察其基因频率的变化及其规律,也要考察基因型频率的变化及其规律。

在共显性遗传和不完全显性的情况下,即各基因型与表型对应的情况下,可依据基因型频率求得基因频率。

假设一个群体共有 N 个个体,在基因 A 和 a 所在的基因座位上,基因型 AA 的个体数为 n_1,其频率为 D;基因型 Aa 的个体数为 n_2,其频率为 H;基因型 aa 的个体数为 n_3,其频率为 R。那么:

$$D=\frac{n_1}{N},\ H=\frac{n_2}{N},\ R=\frac{n_3}{N},\ D+H+R=1。$$

由于每个个体的该基因座位有 2 个基因,因此:

$$p=\frac{2n_1+n_2}{2N}=\frac{n_1}{N}+\frac{1}{2}\times\frac{n_2}{N}=D+\frac{1}{2}H \quad \cdots\cdots ①$$

$$q=\frac{2n_3+n_2}{2N}=\frac{n_3}{N}+\frac{1}{2}\times\frac{n_2}{N}=R+\frac{1}{2}H \quad \cdots\cdots ②$$

如果等位基因有显隐性之分,纯合显性个体与杂合个体在表型上无法区分时,就无法用上述公式计算基因频率;但如果群体中有显隐性之分的等位基因达到了遗传平衡,则可应用遗传平衡定律计算基因频率。

二、遗传平衡定律及其应用

(一)遗传平衡定律

基因频率和基因型频率称为群体的遗传组成,也称为群体的遗传结构。群体所具有的全部遗传信息,称为基因库(gene pool)。不同群体的基因库存在差异,在实际研究中,只能对群体中的某一对基因进行研究,所研究的是狭义基因库。

遗传平衡定律(law of genetic equilibrium),又称 Hardy – Weinberg 定律,是 1908 年英国数学家 Hardy 和德国医生 Weinberg 经过各自对群体遗传的研究获得的一致的结论。即:对于一个能连续随机交配的大的群体来说,如果没有突变、选择、迁移等因素对基因组成和数量的影响,则群体中的基因频率和基因型频率世代相传,保持不变。这样的群体称为遗传平衡群体。

下面举例对遗传平衡定律进行分析。

在前面列举的具有 N 个个体的群体中,假设 N 等于 100 万,n_1、n_2 和 n_3 分别为 60 万、20 万和 20 万,那么,该群体各基因型频率分别是:$D=0.6, H=0.2, R=0.2$。根据公式①和②,可计算出该群体基因 A 和 a 的频率 p 和 q:

$$p=D+\frac{1}{2}H=0.6+\frac{0.2}{2}=0.7$$

$$q=R+\frac{1}{2}H=0.2+\frac{0.2}{2}=0.3$$

根据 Hardy – Weinberg 定律的条件,如果该群体随机婚配,即两性配子随机结合产生子代,亲代产生的配子均为 A 和 a 2 种,频率均为 p 和 q。在无突变、选择、迁移等因素影响的情况下,子 1 代基因型及频率见表 9 – 1。

表 9 – 1　精卵随机结合子 1 代基因型及其频率

			精子			
			A	$p=0.7$	a	$q=0.3$
卵子	A	$p=0.7$	AA	$p^2=0.49$	Aa	$pq=0.21$
	a	$q=0.3$	Aa	$pq=0.21$	aa	$q^2=0.09$

从表 9 – 1 可看出,与亲代相比,子 1 代基因型频率发生了变化:AA 的频率 $D_1=0.49$,Aa 的频率 $H_1=0.21+0.21=0.42$,aa 的频率 $R_1=0.09$。子 1 代基因 A 和 a 的频率没有变化,即:

$$p=D+\frac{1}{2}H=0.49+\frac{0.42}{2}=0.7$$

$$q=R+\frac{1}{2}H=0.09+\frac{0.42}{2}=0.3$$

可以证明,在以后随机婚配所生的各代中都将保持这一遗传组成的平衡状态,即基因频率和基因型频率保持不变。这说明,不平衡的群体只要经过 1 代随机婚配就可达到遗传平衡状态。

上述群体中的个体随机婚配，两性配子随机结合产生子代的频率分布实际上就是一个二项式的展开，即：

$$(p+q)^2=p^2+2pq+q^2=1。$$

这就是遗传平衡公式。式中 p^2、$2pq$、q^2 分别表示遗传平衡状态下 AA、Aa、aa 的基因型频率，即在遗传平衡的群体中，$D=p^2$、$H=2pq$、$R=q^2$。

（二）遗传平衡定律的应用

在研究群体遗传结构时，可根据遗传平衡定律判断群体是否处于遗传平衡状态，更重要的是根据遗传平衡定律计算各种遗传方式下的基因频率。以下主要讨论基因频率的计算，依据遗传平衡定律计算基因频率的前提是将所研究的人类群体近似认为达到了遗传平衡状态。

1. AD 病 例如，已知并指症的群体发病率为 1/2 000，求群体中致病基因频率。在 AD 病，患者绝大多数为杂合子，纯合子患者罕见，可以忽略不计。根据遗传平衡定律，群体达到遗传平衡时，$H=2pq$，由于致病基因 p 频率通常很低，q 近于 1，故 $H\approx 2p$，$p\approx 1/2\times H\approx 1/2\times 1/2\,000=1/4\,000$。

2. AR 病 假设白化病的群体发病率为 10^{-4}，求群体中致病基因的频率及携带者频率。本例中，群体发病率就是纯合子（aa）的基因型频率 R，即达到遗传平衡时，$R=q^2=10^{-4}$。因此，$q=10^{-2}=0.01$，$p=1-q=1-0.01=0.99$；$H=2pq=2\times 0.01\times 0.99\approx 1/50$。

3. X 连锁遗传病 遗传平衡定律同样适用 X 连锁基因。群体中女性 X 染色体上基因频率和基因型频率的分布与常染色体基因相一致，男性则不同；因为男性由于只有 1 条 X 染色体，是等位基因的半合子。男性基因型频率等于基因频率，所以，根据男性的发病率可直接得出基因频率，即男性的表型频率等于相应的基因频率（表 9-2）。

表 9-2 遗传平衡群体中 X 连续基因的基因型及其频率

	女性			男性	
基因型	X^AX^A	X^AX^a	X^aX^a	X^AY	X^aY
基因型频率	p^2	$2pq$	q^2	p	q

例如，在 XR 病，男性红绿色盲的群体发病率为 0.07，即表示群体中红绿色盲致病基因频率 $q=0.07$，而女性色盲患者的发病率 q^2 则为 $0.07^2=0.004\,9$，这与实际观察到的女性红绿色盲发病率 0.05% 是相符的。因为 $q=0.07$，所以，$p=1-q=1-0.07=0.93$，女性携带者频率 $2pq=2\times 0.93\times 0.07=0.130\,2\approx 13\%$。这说明女性发病率虽然很低，但携带者频率却高达 13%。

再如，在 XD 病，女性患者有 2 种基因型，即 $X^AX^A(p^2)$ 和 $X^AX^a(2pq)$，而男性患者只有 1 种基因型 $X^AY(p)$，可以看出 XD 病中男性发病率就等于致病基因频率。当罕见的 XD 病的致病基因频率 p 很低时，可以忽略不计，因此，男性发病率与女性发病率的比例为：

$$\frac{p}{2pq+p^2}=\frac{p}{p(2q+p)}=\frac{1}{2(1-p)+p}=\frac{1}{2-p}\approx\frac{1}{2}$$

即女性发病率是男性发病率的 2 倍。已知遗传性肾炎男性发病率为 1/10 000，这说明 $p=1/10\,000$，女性发病率 $=2\times 1/10\,000=2/10\,000$。

第二节 影响群体遗传平衡的因素

遗传平衡定律所要求的群体是理想群体，严格地说，这样的群体在自然界中是不存在的，人类更是如此。如果群体遗传平衡所要求的条件不能满足，群体的遗传结构将发生改变，群体则处于非遗传平衡状态。从而使遗传平衡遭到破坏。下面讨论突变、选择、迁移、随机遗传漂变等因素对群体遗传平衡的影响。

一、突变

突变(mutation)是影响遗传平衡的重要因素。突变改变群体的遗传组成，即改变群体的基因频率和基因型频率。基因突变是双向的，包括正向突变和回复突变。通常回复突变率只有正向突变率的1/10。突变率通常很低，一般用每代、每1 000 000个基因中基因突变的次数(n)来表示，即 $n\times10^{-6}$/(基因·代)。

在一对等位基因A(p)和a(q)中，设A→a的突变率为μ，a→A的回复突变率为ν，每一代中A→a比率为$p\mu=(1-q)\mu$，a→A的比率为$q\nu$。若$(1-q)\mu>q\nu$，则q增加；若$(1-q)<q\nu$，则p增加；若$(1-q)\mu=q\nu$，则p、q保持不变，群体达到遗传平衡。

对$p\mu=(1-q)\mu=q\nu$进行变换，则有：

$$\mu-q\mu=q\nu$$

$$\mu=q\nu+q\mu=q(\mu+v)$$

$$q=\mu/(\mu+\nu),$$

$$p=v/(\mu+v)$$

这说明在没有选择作用的情况下，群体的基因频率完全由μ和v决定，即群体遗传平衡是由等位基因的双向突变维持的。群体中的某些既无害、也无益的中性突变就是以这样方式维持群体遗传平衡的。

二、选择

选择(selection)包括自然选择(natural selection)和人工选择(artificial selection)，通常所说的选择是指自然选择。选择是指由于基因型的差别导致生存能力和生育能力的差别。简单地说，选择就是适者生存，不适者淘汰。生物包括人类具有高度的个体独特性，每个个体都是独一无二的，都产生于同一基因库。选择就是对突变产生的变异进行整理的过程。在具有个体特异性的生物群体中，具有更适应当时生态环境的基因组的个体，比群体中的其他个体成员存活的可能性更大。这样群体中适应性强的基因型个体增加，不适应的个体逐渐减少乃至消失。

环境条件是不断变化的，生物体的适应性是相对的，选择不断地使群体的遗传组成做出相应的变化，从而使群体中的个体建立新的适应关系。也可以说，群体的遗传组成处于动态平衡状态。每个基因都控制着生物体的生理特性和形态结构，这些生理特性和形态结构又或多或少地影响着个体的生活力和繁殖力。绝大多数基因会受到选择的作用，使适应能力差的基因型逐渐减少，适应能力高的基因型逐代增多，这样就直接改变了群体的

遗传组成。

选择的作用在于增加或减少群体中个体的适合度(fitness,f)。由于大多数情况下,选择的作用都是减少个体的适合度,当适合度为0时,个体被淘汰。因此,通常所说的"被选择"即指被淘汰。适合度是指在特定的环境中,某一基因型的个体能生存并生育后代的能力,通常用同一环境中不同个体的相对生育率来衡量。选择系数(selection coefficient, S)是指在选择作用下降低的适合度,是选择强度即选择压(selection press)的度量,它表示某一基因型在群体中不利于生存的程度。适合度和选择系数的关系是:$S=1-f$,$f=1-S$。

例如,根据丹麦的一项调查,108名软骨发育不全性侏儒共生育了27个孩子,而这些患者的457个正常同胞共生育了82个孩子。如以正常人的生育率为1,软骨发育不全性侏儒患者的相对生育率 $f=(27/108)/(82/457)=0.20$,选择系数 $S=1-f=1-0.2=0.8$。适合度大于1,表示该病患者比正常人生存和生殖能力强,如镰形细胞贫血症杂合子个体抗恶性疟能比正常人强,其适合度为1.26(表9-3)。

表9-3 几种遗传病患者的适合度

遗传病	适合度
Tay-Sachs病	0
视网膜母细胞瘤	0
软骨发育不全	0.2
甲型血友病	0.29
多发性神经纤维瘤	0.41(男),0.75(女)
亨廷顿舞蹈病	0.82(男),1.25(女)
镰形细胞贫血症(恶性疟高发区)	0(患者),1.26(携带者)

不同基因型的个体适合度不同,而选择对不同基因型起的作用不同,即有各自的选择系数。下面分别讨论不同情况下的选择作用。

(一)选择对常染色体显性基因的作用

对一种有害的显性突变,如果这种突变会引起个体死亡或不能生育后代,即适合度为0,那么选择是有效的,只需1代即可淘汰有害基因;如果这种突变是有害的,但不是致死的,那么患病个体的生育力将降低,如前述的软骨发育不全性侏儒。

经1代选择后,显性基因A的频率将降低。在选择的作用下,每一代被淘汰的基因A为Sp。

因为 $p=D+(1/2)H$,纯合子患者罕见,可忽略不计,所以,$p=(1/2)H$,被淘汰的基因将由突变基因来补充,以维持平衡,因此,显性基因突变率 $v=Sp=S(1/2)H$。

例如,在丹麦的某医院中,几年来共出生97 075名婴儿,其中,有10名婴儿患软骨发育不全性侏儒(AD),发病率 $H=10/94\,075$。已知该病的适合度 $f=0.20$,选择系数 $S=0.80$,因此,该病的致病基因突变率:

$$
\begin{aligned}
v &= Sp = S(1/2)H \\
&= 0.80\times 0.5\times 10/94\,075 \\
&= 43\times 10^{-6}/(\text{基因·代})
\end{aligned}
$$

(二)选择对常染色体隐性基因的作用

与显性突变相比,选择对隐性突变基因的作用是有限的,只有隐性纯合子患者面临选择,而杂合子不被选择。即使隐性纯合个体的适合度为0,即突变致死,也仅淘汰少数隐性基因。被淘汰的隐性基因(Sq^2)要由新的突变基因来补充,以使群体保持遗传平衡,因此,隐性基因突变率 $u = Sq^2$。

例如,已知AR病苯丙酮尿症的群体发病率为1/16 500≈0.000 06,$f = 0.30$, $S = 0.70$,因此,该病致病基因突变率 $u = Sq^2 = 0.70 \times 0.00006 = 42 \times 10^{-6}$/(基因·代)。

(三)选择对杂合子的影响

选择作用对杂合体的影响不同,并不都使被选择的基因的频率下降。例如,母亲血型是Rh^-,父亲为Rh^+,则胎儿的血型是Rh^+,这种情况下母亲与胎儿之间血型不相容,会引起新生儿溶血,即可对新生儿形成一种负向选择,使胎儿溶血死亡。在恶性疟高发区,存在镰形细胞贫血症正向选择的情况。在非洲某些地区,镰形细胞贫血症致病基因 β^S 频率较高,目前并没有证据证明这样高的频率是由较高的基因突变造成的。现在了解到在这些地区,基因型为 $\beta^A\beta^S$ 的杂合个体比 $\beta^S\beta^S$(正常人)和 $\beta^A\beta^A$ 纯合子(患者)有更高的适合度,这称为杂合子优势。这是由于环境因素对杂合子的正向选择所致,即因为这些地区是恶性疟高发区,杂合体对恶性疟有更高的抵抗力,使杂合体表现出比正常纯合子有更强的生存能力,能生育更多的后代,其机制可能是发生镰变的红细胞不利于疟原虫寄生。

三、迁移

迁移(migration)是指具有一定基因频率的群体中一部分个体迁入基因频率不同的另个群体并杂交定居。迁移改变了迁出群体和迁入群体的基因频率,即基因从一个群体扩散到另一群体。因此,迁移也称基因流(gene flow)。人类不同地区、不同种族的基因频率存在差异,大规模迁移引起基因频率改变,形成巨大的迁移压力,从而改变群体的遗传结构。

例如,欧洲和西亚白种人中,苯硫脲(PTC)的尝味能力缺乏者(味盲,基因型 tt)频率为36%,味盲基因(t)频率为$\sqrt{36\%} = 0.60$。我国汉族人群PTC味盲者频率为9%,味盲基因(t)频率为$\sqrt{9\%} = 0.30$。而聚居在两者之间的我国宁夏一带的回族人群,PTC味盲者频率为20%,味盲基因(t)频率为$\sqrt{20\%} = 0.45$。究其原因,可能是在唐代,欧洲人和西亚人,尤其是古代波斯人沿丝绸之路到长安进行贸易活动,以后又在宁夏附近定居,与汉族人通婚形成基因流所致。

四、随机遗传漂变

突变、选择、迁移等因素对群体遗传平衡的影响,都是基于一个相当大的、环境条件不变的、平衡的群体。对于生物个体数量有限的、较小的群体,其遗传组成的变化即基因频率和基因型频率的改变,与突变、选择、迁移等因素对群体的遗传组成的影响不同。这种小群体的遗传组成的改变存在随机遗传漂变现象。由于群体较小和偶然事件造成群体中基因频率的随机波动,称为随机遗传漂变(random genetic drift)。

人类不同群体的ABO血型系统的基因频率存在差异,这种差异不是选择的结果,而是主要是随机遗传漂变所致。例如,北美印第安人 I^A 频率为0.18,I^B 频率为0.009,i 频

率为0.973;O血型者占94.6%,A血型者占3.6%,B血型者占1.8%,AB血型者为0。但在蒙大拿州的勃拉克斐特(Blakfeet)印第安人中,I^A频率高达0.5,远高于任何印第安人的群体,这可能是由于在小群体中的随机遗传漂变所致。

遗传漂变在小群体中特别有效,它可以掩盖甚至违背选择所起的作用,即使选择不利的等位基因,只要它不是致死的,也会因漂变而建立和固定;相反,选择有利的等位基因,在选择作用还没有充分表现出它的效应之前,有可能因漂变而淘汰。

由于环境的激烈变化,随机漂变使群体中的个体数急剧减少,甚至面临灭绝,但濒临灭绝的等位基因频率由于偶然发生变异,就会逃过淘汰选择而存活,这类似于群体通过了生存瓶颈(bottle neck)。一个大群体通过瓶颈后,可由剩下的少数个体再扩展成原来的规模群体,这种群体数量的消长对遗传组成造成的影响称为瓶颈效应(bottle neck effect)。这是一种极端的遗传漂变作用。群体经过这种"瓶颈"的漂变作用后,所造成的等位基因频率的变动将较长时期地保留在后代群体中。

生物在湖泊、隔离的森林和其他的隔离生境中,还存在着一种极端的遗传漂变作用,即建立者效应(founder effect):由少数个体的基因频率决定了他们后代中基因频率的效应。在海岛上生活的许多物种的群体,他们现在可能有几十万、或者几百万个个体组成,但是他们却是很久以前到来的少数 "入侵者"的后裔。

人类群体中也存在奠基者效应。例如,卟啉症是一种罕见的AD病,但南非居民发病率却高达1%,据回顾性调查发现,这些南非居民的共同祖先是17世纪移居到好望角的一对荷兰夫妇。再如,美州印第安人缺乏B型血型,这些印第安人的祖先是在大约2万年前从亚州迁去的,很可能最初的群体中就不含B型等位基因,或在最初的几个世纪因随机漂变导致该等位基因从群体中消失所致。

第三节 近婚系数

按照遗传平衡的要求,在一个无限大的群体中,随机婚配才能保证群体遗传结构的相对稳定,但实际生活中,人们的活动范围往往受到限制,婚配也受到地域、习俗、经济状况、兴趣爱好等因素的影响,不可能是完全随机的,并且还存在有亲缘关系的人群间相互婚配的现象,这样将影响群体的遗传平衡;并且有亲缘关系的人婚配往往会使隐性致病基因形成纯合子的机会增多,导致隐性遗传病发病率增高。近亲婚配的危害性可通过亲缘系数来评价,也可通过近婚系数来衡量。前者在单基因遗传病一章已作介绍,本节主要讨论近婚系数。

近婚系数(coefficient of inbreeding, *F*)是指近亲婚配的配偶从共同祖先得到同一基因,又同时将这一基因传递给他(她)们的子女,使之成为纯合子的概率。

一、常染色体基因的近婚系数

(一)同胞兄妹婚配

假设某同胞兄妹的父母(P_1、P_2)的某一基因座位上的等位基因分别为A_1A_2、A_3A_4(图9-1),任何一个等位基因传递给子代的概率都是1/2,P_1分别通过B_1和B_2共4步将A_1传

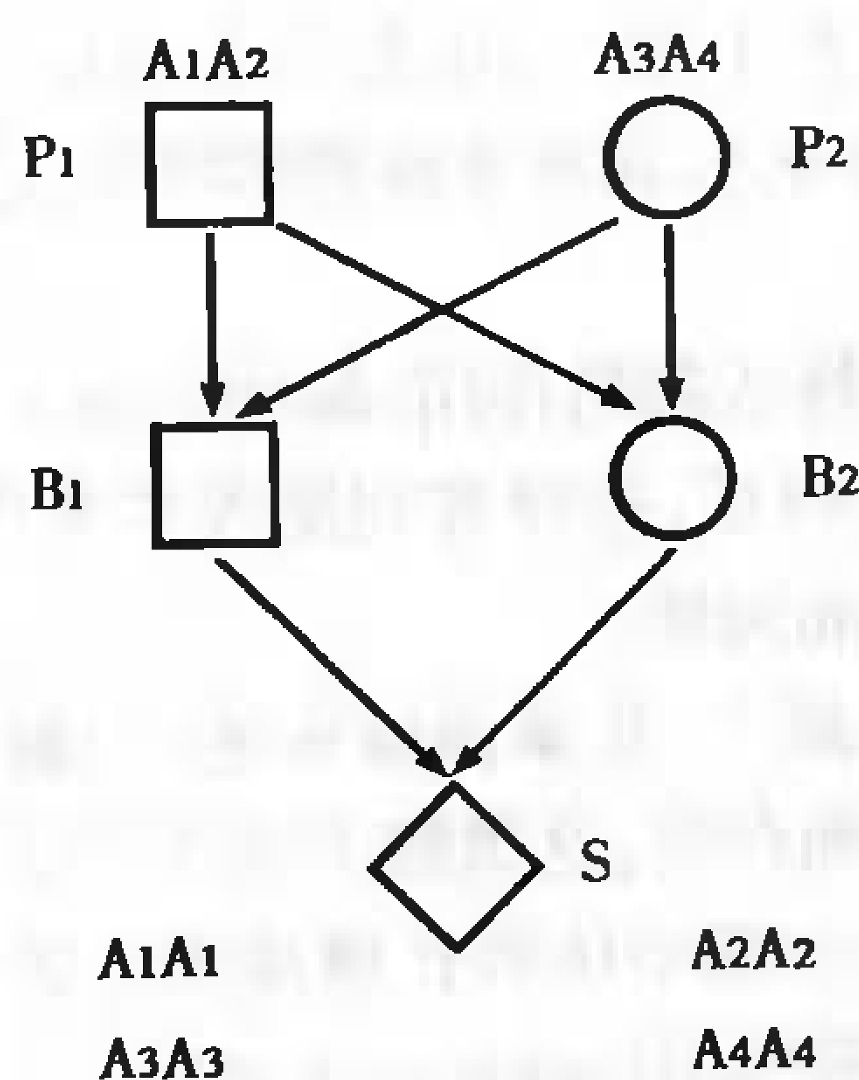


图 9－1　同胞兄妹婚配基因传递

递给 S 并使之纯合，所以，来自 P_1 的 A_1 基因传给 S 形成 A_1A_1 纯合子的概率为 $(1/2)^4$。同理，S 形成纯合子 A_2A_2、A_3A_3、A_4A_4 的概率也都是 $(1/2)^4$。如果只考虑纯合子，而不考虑是来自父亲的 A_1 或 A_2 的纯合，还是来自母亲的 A_3 或 A_4 的纯合，即 4 种纯合子中任一种出现的概率，也就是总概率为 $4\times(1/2)^4=1/4$。这就是同胞兄妹婚配的近婚系数，即：$F=1/4$。同胞兄妹为一级亲属，父子之间、母女之间也是一级亲属。一级亲属的近婚系数均为 1/4。

（二）舅甥女婚配

P_1 把基因 A_1 通过 B_1 经过 2 步传递给 S，通过 B_2 经过 3 步传递给 S。这样基因 A_1 共经过 5 步传递给 S 并纯合为 A_1A_1（图 9－2）。基因 A_2、A_3 和 A_4 同样也需 5 步传递给 S 并纯合为 A_2A_2、A_3A_3 和 A_4A_4，其概率各为 $(1/2)^5$。因此，$F=4\times(1/2)^5=1/8$。舅甥女为二级亲属，即二级亲属的近婚系数为 1/8，近婚系数为 1/8 的二级亲属还包括姑侄、叔侄女、姨外甥等。

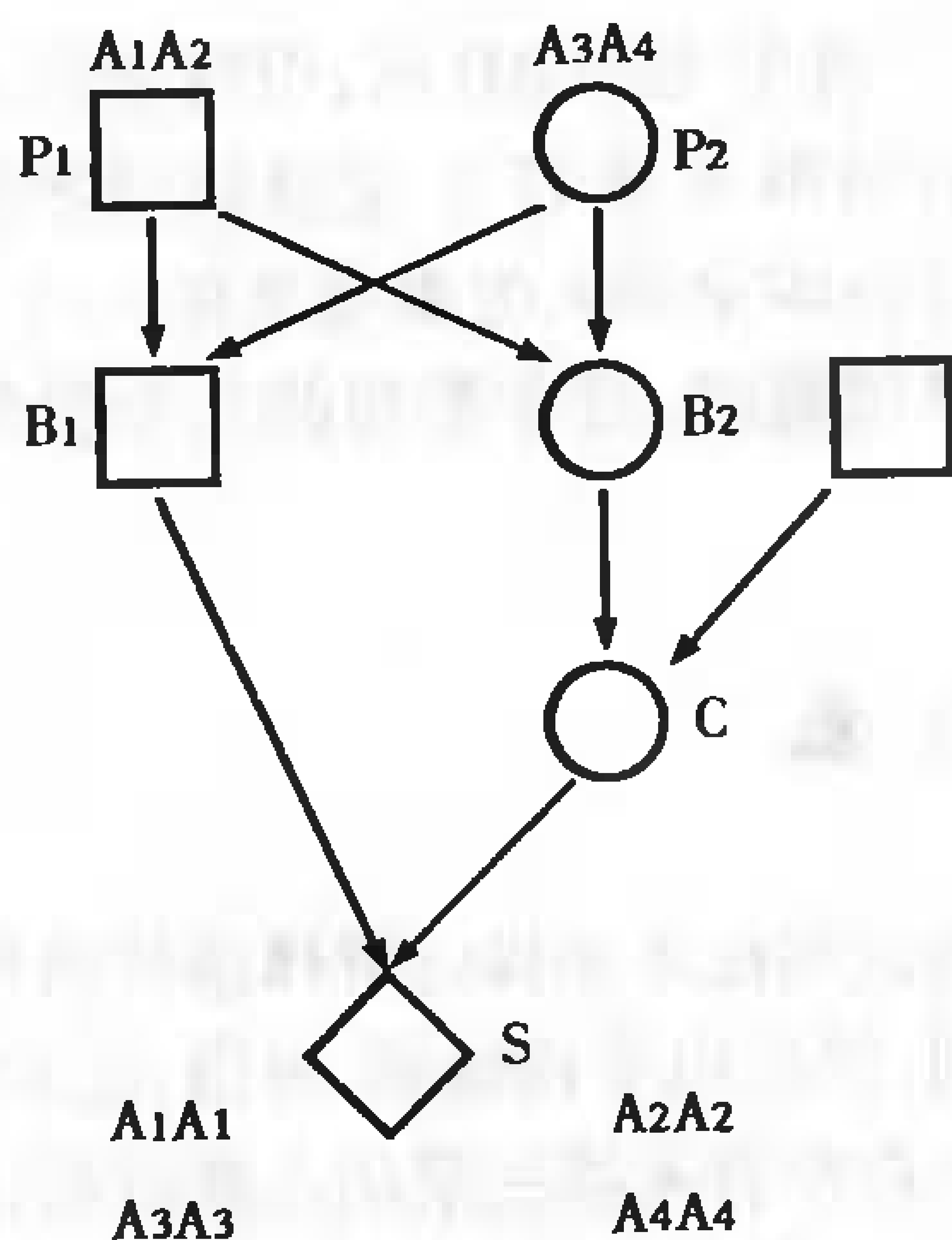


图 9－2　舅甥女婚配基因传递

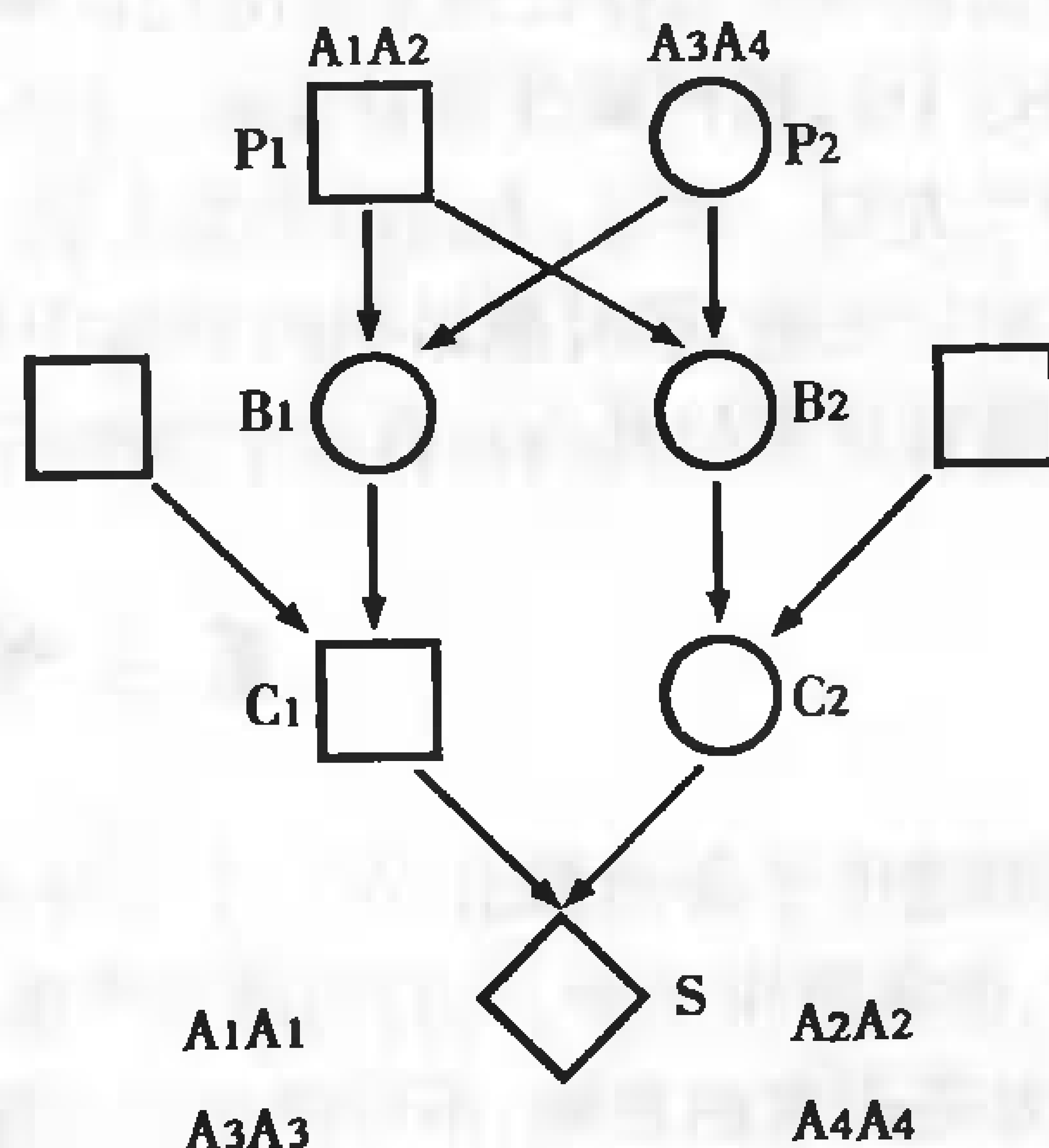


图 9－3　表兄妹婚配基因传递

（三）表兄妹婚配

表兄妹有姨表兄妹、姑表兄妹、舅表兄妹之分。如果表妹是姑所生称姑表兄妹，表妹是舅所生称舅表兄妹。图 9－3 为姨表兄妹婚配基因传递图。表兄妹及堂兄妹为三级亲属，三级亲属间婚配，共同祖先的任一基因需经 6 步传递到 S 并使之纯合，因此，$F=(1/2)^6$，即三级亲属的近婚系数为 1/16。

（四）二级表兄妹婚配

与表兄妹婚配相比，二级表兄妹婚配的基因传递增加 2 步，即 S 成为纯合子需 8 步基因传递（图 9－4），所以，二级表兄妹的近婚系数为 $4\times(1/2)^8=1/64$。

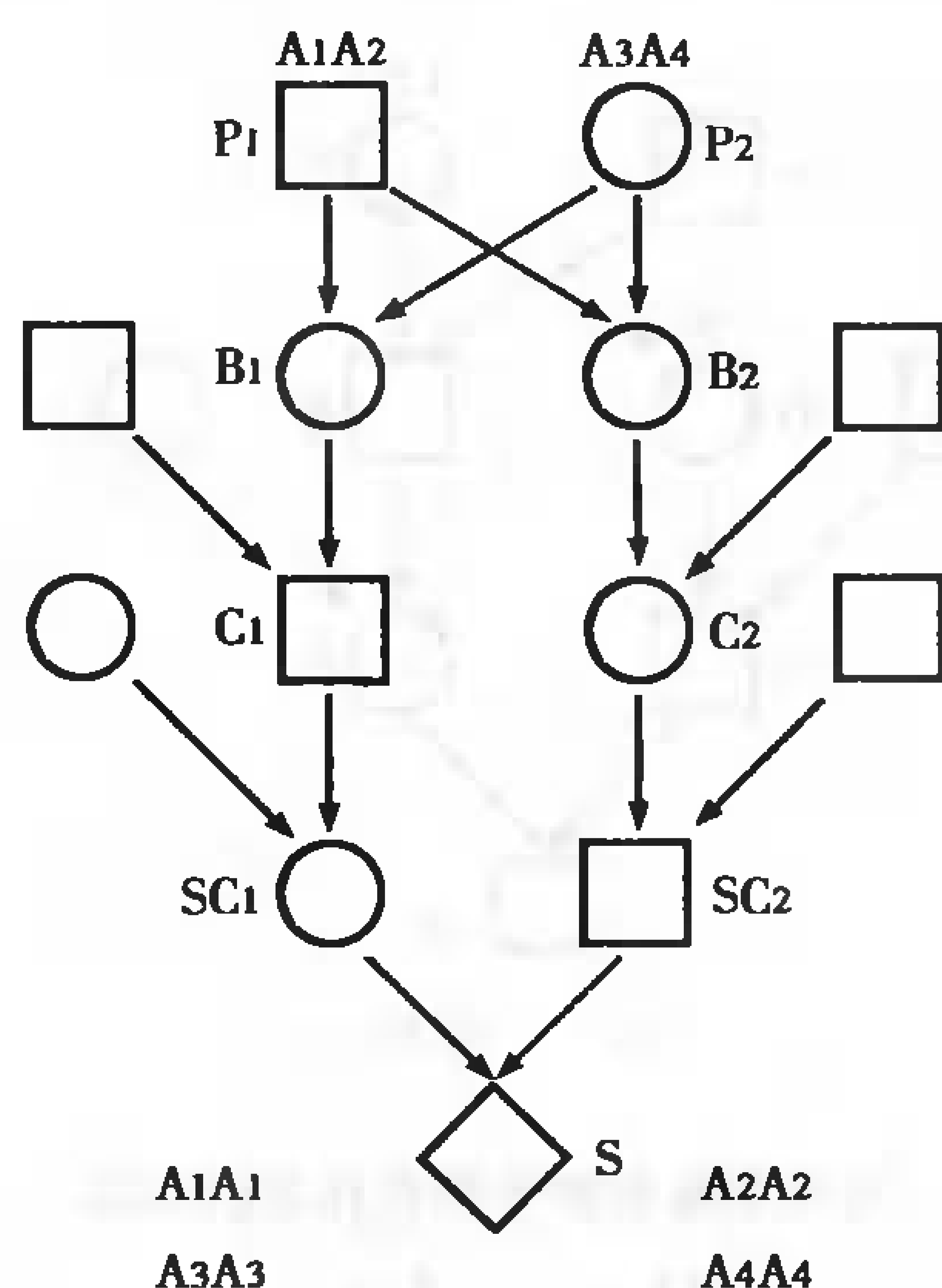


图 9-4 二级表兄妹婚配基因传递

(五)半同胞兄妹婚配

同父异母或同母异父的兄妹为半同胞兄妹,他(她)们只有一个共同祖先。因此,半同胞兄妹的近婚系数为 $2\times(1/2)^4=1/8$(图 9-5)。

二、X 染色体基因的近婚系数

X 染色体上的基因传递与常染色体不同:①女性有 2 条 X 染色体,基因可以纯合;男性只有 1 条 X 染色体,不可能形成纯合子。因此,只计算生女儿情况下的近婚系数;②女性 X 染色体上基因向后代传递的规律与常染色体相同,男性 X 染色体上基因传递给女儿的概率为 1,传给儿子的概率为 0。

(一)姨表兄妹婚配

P_1 将 X_1 传给 B_2 的概率为 1,B_1 传给 C_1 的概率是 1/2,C_1 传递给 S 的概率又是 1;P_1 将 X_1 传递给 B_2 的概率是 1,B_2 传给 C_2、C_2 再传给 S 的概率均为1/2,所以 P_1 将 X_1 传给 S 并使之纯合的概率为:$1^3\times(1/2)^3=(1/2)^3$。同理,P_2 的 X_2、X_3 传给 S 并使之纯合的概率都是 $1\times(1/2)^5=(1/2)^5$(图 9-6)。因此,姨表兄妹的 X 染色体上基因的近婚系数 $F=(1/2)^3+(1/2)^5=3/16$。

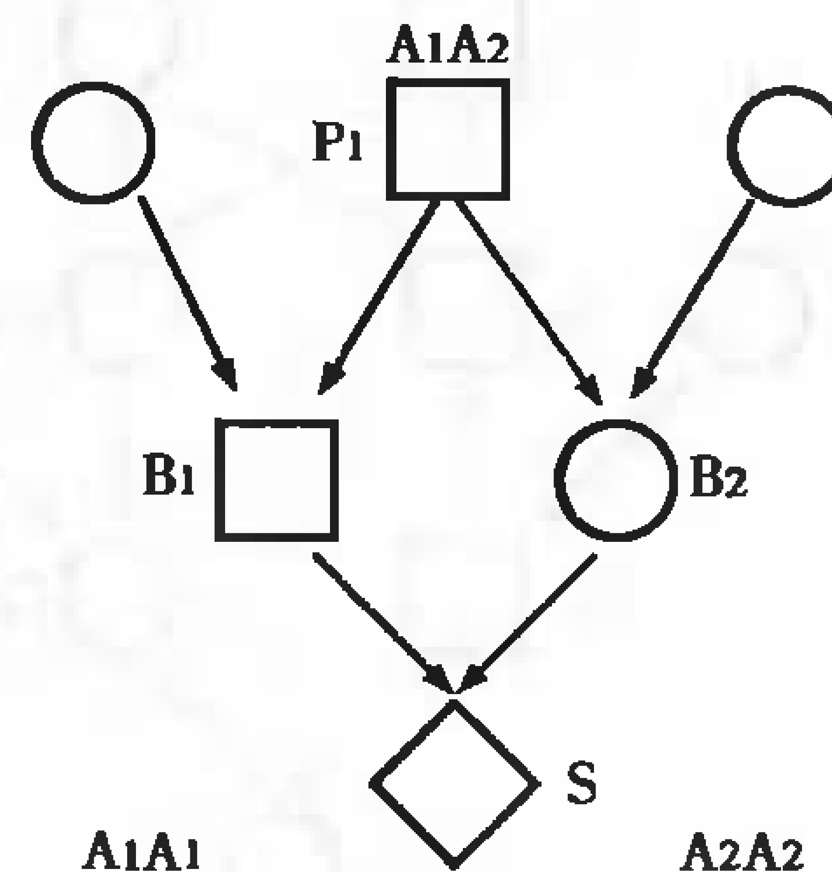


图 9-5 半同胞兄妹婚配基因传递

(二)舅表兄妹婚配

P_1 的 X_1 只能通过 B_1 传递给 S,不能通过 B_2 传递给 S,所以 X_1 不能在 S 纯合;P_2 的 X_2、X_3 需 6 步传递到 S 并纯合,其中 2 步的概率为 1,其余 4 步的概率均为 1/2(图 9-7),所以,舅表兄妹的 X 染色体上基因的近婚系数 $F=0+1^2\times(1/2)^4=1/8$。

(三)姑表兄妹婚配及堂兄妹婚配

在这两种婚配方式下,不管是 P_1 的 X_1,还是 P_2 的 X_2、X_3,都不能在 S 纯合,所以,姑表兄妹及堂兄妹婚配,X 染色体上基因的近婚系数 $F=0$(图 9-8、9)。

以上的计算分析表明,同是三级亲属,常染色体上基因的近婚系数均为 1/16,而 X 染色体上基因的近婚系数则不同,危害性最大的是姨表兄妹婚配,近婚系数高达 3/16。

三、平均近婚系数

近婚系数是从家庭的角度说明近亲婚配的危害性,而评价群体近亲婚配的程度及其危害性常用平均近婚系数(average coefficient of inbreeding, a),平均近婚系数通过以下公式计算:

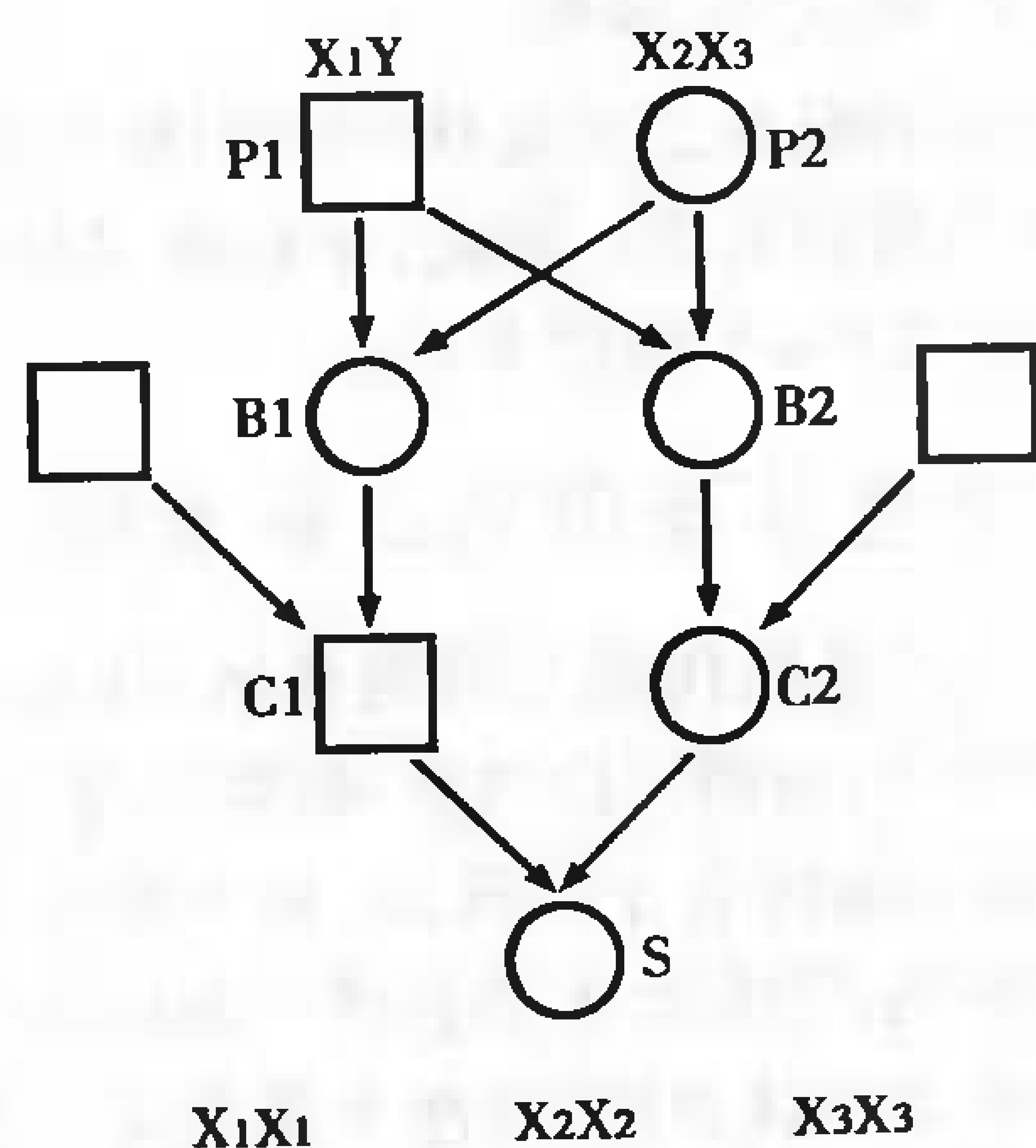


图 9-6　姨表兄妹婚配 X 染色体基因传递

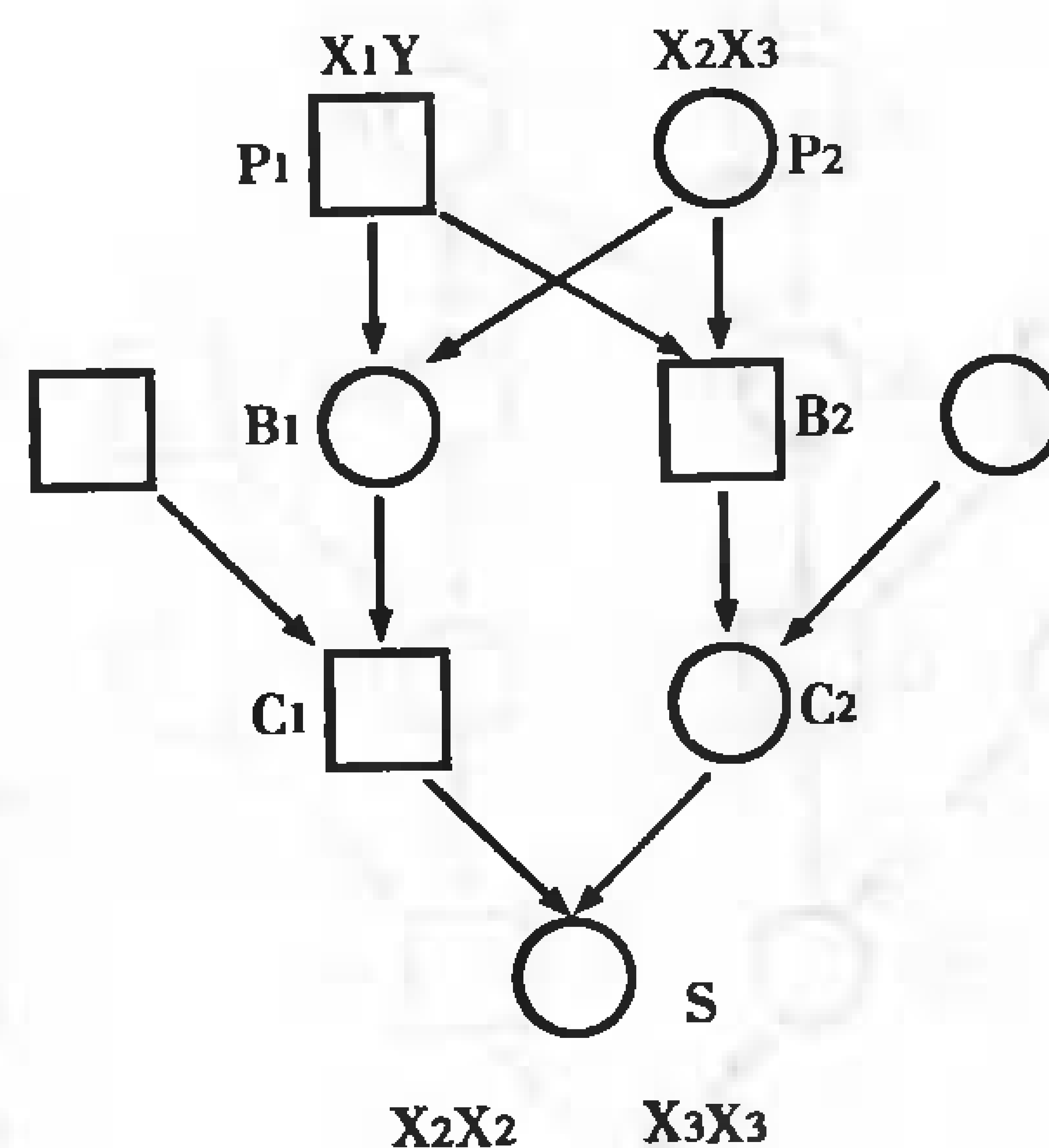


图 9-7　舅表兄妹婚配 X 染色体基因传递

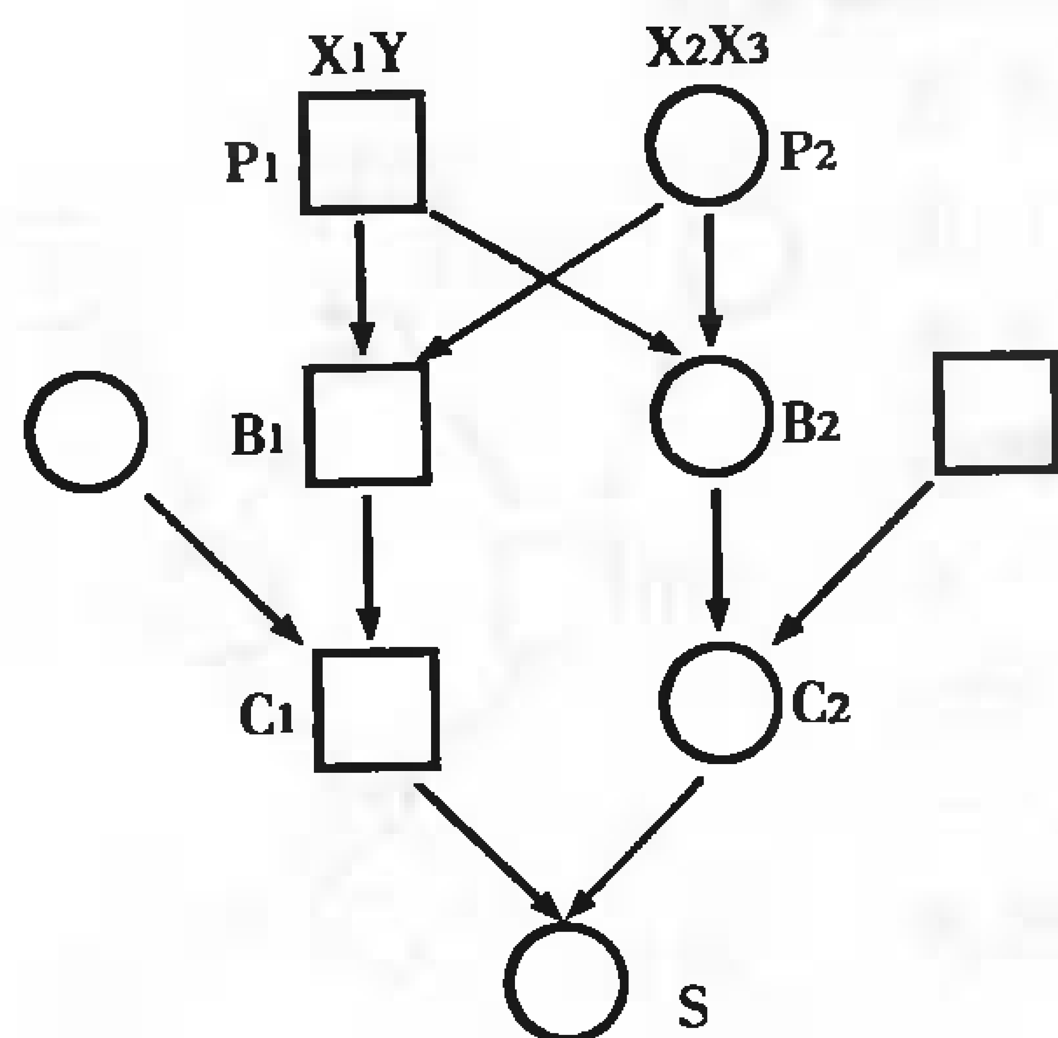


图 9-8　姑表兄妹婚配 X 染色体基因传递

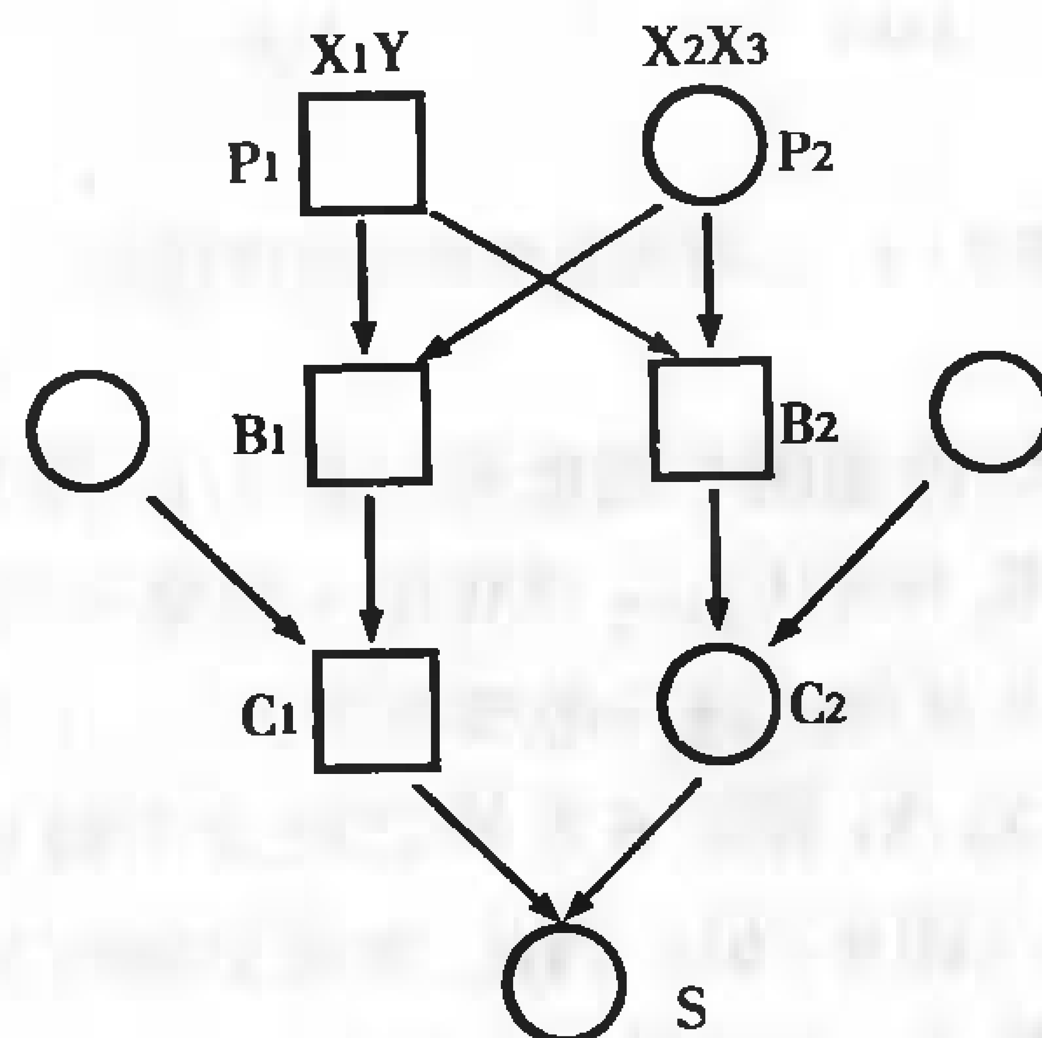


图 9-9　堂兄妹婚配 X 染色体基因传递

$$a = \frac{\sum M_i F_i}{N}$$

公式中 M_i 表示某种类型的近亲婚配对数，F_i 表示相应近亲婚配的近婚系数，N 表示婚配总对数。例如，在一个群体的 1 000 例婚配中，兄妹婚 1 例，舅甥女婚 2 例，表兄妹婚 45 例，二级表兄妹婚 18 例，二级半同胞兄妹婚 6 例，三级表兄妹婚 3 例，其余为非亲缘关系婚配。这一群体平均近婚系数(a)计算如下：

$$a = \frac{\sum M_i F_i}{N}$$

$$= \frac{1\times\frac{1}{4}+2\times\frac{1}{8}+45\times\frac{1}{16}+18\times\frac{1}{64}+6\times\frac{1}{32}+3\times\frac{1}{256}}{1\,000}$$

$$= 0.003\,79$$

一般来说，a 值达到 0.01 就算相当高了。在山区和某些少数民族地区，由于地理的阻隔及婚配习俗的影响，近亲婚配的现象较普遍，a 值往往较高。

第四节 遗传负荷

遗传负荷(genetic load)是指一个群体由于有害基因或致死基因的存在而使群体适合度降低的现象。一般用群体中平均每个个体携带的有害基因数量来表示。

根据有害基因或致死基因的来源,遗传负荷分为突变负荷和分离负荷 2 种。突变负荷(mutation load)是指基因突变产生有害基因或致死基因,使群体适合度下降而给群体带来的遗传负荷。近亲婚配导致分离负荷(segregation load),分离负荷是指适合度较高的杂合子由于基因分离而产生适合度较低的纯合子,降低群体适合度而给群体带来的遗传负荷。

有害基因或致死基因的存在给群体的遗传组成造成压力,不同群体遗传负荷不同。一级亲属婚配后代发病调查是估计群体遗传负荷方法的一种方法(表 9-4)。

表 9-4 一级亲属(同胞)婚配所生子女 AR 病的发病情况

国家	调查年份	一级亲属婚配所生子女数	严重缺陷个体数	严重智力低下个体数
美国	1967	18	6	—
英国	1967	13	8	—
捷克	1971	161(存活 138)	60	40
加拿大	1982	21	9	—
合计		190	83	40
发病率		—	44%	21%

具体估算方法是,假设一个人是某种 AR 病基因的肯定携带者,那么其同胞有 1/2 可能携带该基因。他(她)们婚配后其子女患 AR 病的风险为:$1\times1/2\times1/4=1/8$。如果该个体携带 n 个 AR 病基因,那么这个人与其同胞婚配,其子女患 AR 病的风险则为:$n\times1/2\times1/4=n/8$。

根据表 9-4 可知,一级亲属婚配子女 AR 发病率为 44%,即 $n/8=44\%$,因此,$n=3.52$;如果将严重智力低下的情况考虑进去,则 $n/8=65\%$,那么,$n=5.2$。这说明,在以上所调查的 4 个国家的人群中,一个人可能携带 4~5 个有害基因。值得注意的是,由于一级亲属婚配的实例较少,没有更多的资料可供研究,所以在比较群体遗传负荷时,要注意数据来源的年代、调查的实际例数及数据来源的群体大小。根据目前的调查,美国人的遗传负荷为 5~8,日本人的遗传负荷为 4~5,我国人群的遗传负荷为 5~6。

小 结

群体是指同一物种生活在某些地区并能够互相交配的个体群。群体遗传学是研究生物群体的遗传结构及其变化规律的遗传学分支学科,它应用数学和统计学方法来研究群体中基因频率和基因型频率以及两者之间的关系,同时研究突变、选择、迁移、随机遗传漂变、隔离等因素与群体遗传结构的关系。

群体的遗传结构又称为群体的遗传组成,是指群体的基因种类和频率、基因型的种类

和频率。基因频率是指群体中某一基因占该基因座位上全部等位基因的比率。基因型频率是指某一基因型的所有个体在所研究群体中所占的比率。

如果某一群体是一个能连续随机交配的、大的群体，且没有选择、突变、迁移、随机遗传漂变等因素对基因组成和数量的影响，该群体即处于遗传平衡状态，其基因频率和基因型频率将世世代代保持不变。此称为群体遗传平衡定律。

遗传平衡定律的公式是：$(p+q)^2=p^2+2pq+q^2$，$p+q=1$。式中 p^2、$2pq$、q^2 分别表示遗传平衡状态下 AA、Aa、aa 的基因型频率，即在遗传平衡的群体中，$D=p^2$、$H=2pq$、$R=q^2$。

遗传平衡定律所要求的条件实际上就是影响遗传平衡的因素，这些条件如不具备，群体的遗传结构将发生改变。

近婚系数是指近亲婚配的配偶从共同祖先得到同一基因，又同时将这一基因传递给他(她)们的子女，使之成为纯合子的概率。评价群体近亲婚配的程度及其危害性常用平均近婚系数，平均近婚系数通过公式 $a=\Sigma M_i F_i/N$ 计算。

遗传负荷是指一个群体由于有害基因或致死基因的存在而使群体适合度降低的现象。一般用群体中平均每个个体携带的有害基因数量来表示。根据有害基因或致死基因的来源，遗传负荷分为突变负荷和分离负荷 2 种。

研究群体遗传学的基本原理对了解致病基因在人群中的分布及其变化规律，运用这些规律制定遗传病的预防措施，提高人类遗传素质具有积极意义。

（李学英）

第十章　免疫遗传学

免疫遗传学是免疫学与遗传学相互渗透形成的边缘学科，它研究机体免疫反应的遗传基础及遗传调控，从分子水平阐明人类免疫现象的遗传和变异规律以及与遗传有关的免疫性疾病的遗传背景，揭示免疫现象的本质，从而为疾病的预防、诊断、预后判断及治疗提供帮助。因此，免疫遗传学的研究正日益受到重视。

第一节　血型的遗传

自1900年奥地利医生 Landsteiner 首先发现了 ABO 血型起，迄今又有 MN、P 和 Rh 等约20余种血型系统被陆续发现，其中与临床关系最为密切的为 ABO 血型系统和 Rh 血型系统。

一、ABO 血型系统

ABO 血型系统是输血和器官移植中非常重要的血型系统。ABO 血型抗原除分布于红细胞膜外，还分布于淋巴细胞、血小板、内皮细胞及上皮细胞等细胞膜上，故称之为组织血型抗原(histo – blood group)。

ABO 抗原由3组基因编码，即，位于9q34上的 I^A、I^B 和 i 复等位基因，同时位于19号染色体的 H、h 及 Se、se 等位基因。其中，I^A、I^B 为显性基因，i 为隐性基因；H 为显性基因，h 为无效基因；Se 为显性基因，se 为无效基因。

H 基因的产物是 L – 岩藻糖转移酶，该酶将 L – 岩藻糖连接在细胞膜的 H 前体物质(四糖体)上形成 H 物质，具有 H 抗原特异性；几乎所有个体(HH 或 Hh)的细胞膜上都有 H 抗原，H 抗原是 A、B 抗原的前体。I^A 基因的产物是乙酰半乳糖胺转移酶，该酶将乙酰半乳糖胺转移至 H 物质的半乳糖末端形成 A 抗原，故基因型 I^AI^A 和 I^Ai 的个体表现为 A 血型；I^B 基因的产物是半乳糖转移酶，该酶将半乳糖转移至 H 物质的半乳糖末端故形成 B 抗原，故基因型 I^BI^B 和 I^Bi 的个体表现为 B 血型；基因型 I^AI^B 的个体能产生上述2种酶，能将乙酰半乳糖胺和半乳糖转移至 H 物质上，形成含 A 抗原和 B 抗原的复合抗原，表现为 AB 血型；基因型 ii 的个体因无上述2种酶的产生，细胞膜上只有 H 抗原，表现为 O 血型(图10 – 1、2)。

在人血清中存在红细胞本身所不具备抗原的天然抗体：A 血型者血清中存在抗 B 抗体，B 血型者血清中存在抗 A 抗体，O 血型者血清中既有抗 A 抗体又有抗 B 抗体，AB 血型者血清中2种抗体均不存在。

极少数基因型为 hh 的个体，由于没有 H 基因，细胞膜上无 H 抗原，也不能形成 A 抗原、B 抗原，此种 O 血型者与细胞膜上有 H 抗原的 O 血型者不同，称为 Oh 血型。因 Oh 血型在印度孟买被首次发现，故又称孟买型。孟买型者血清中有抗 A 抗 B 与抗 H，所以除了

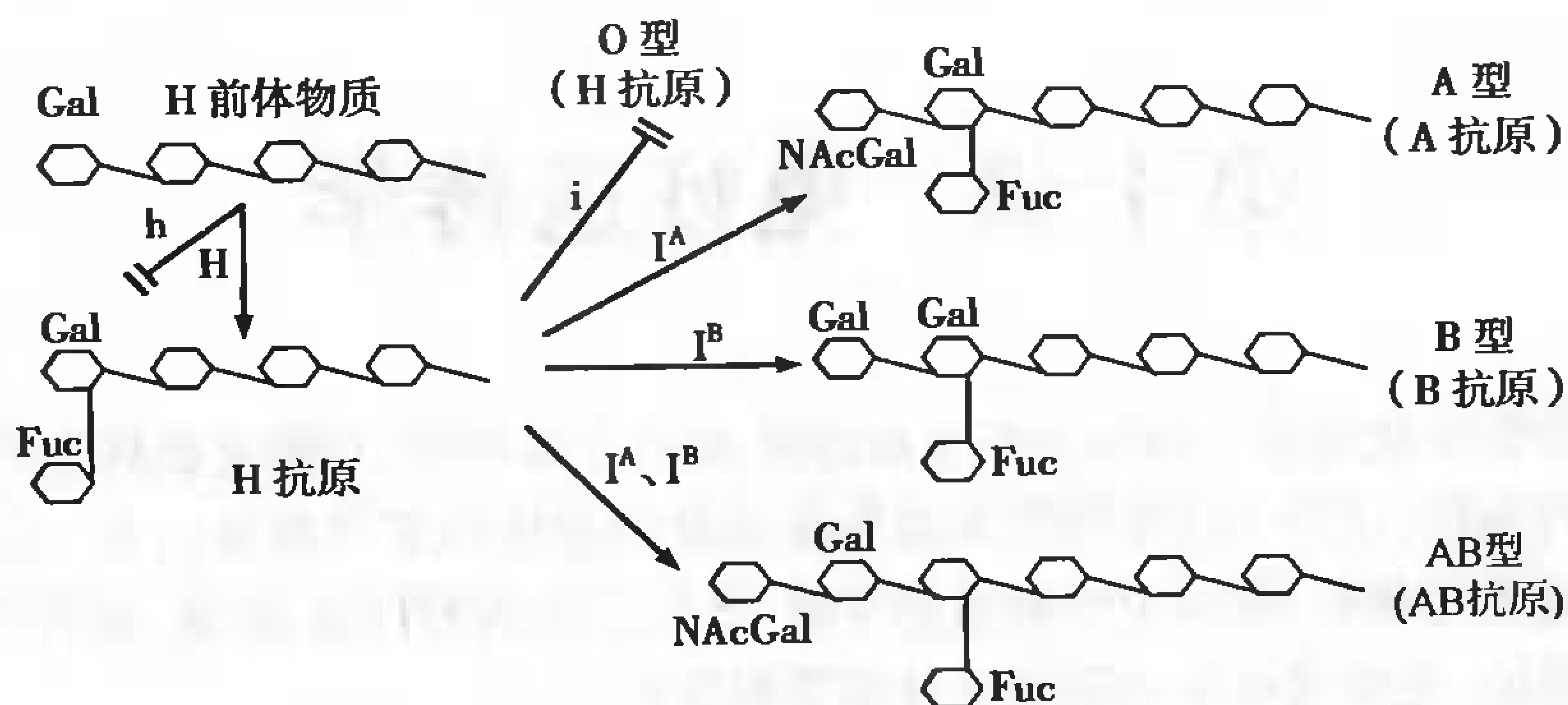


图 10－1　ABO 血型系统血型抗原形成的遗传机制
Gal：半乳糖；Fuc：岩藻糖；NAcGal：*N*－乙酰半乳糖胺

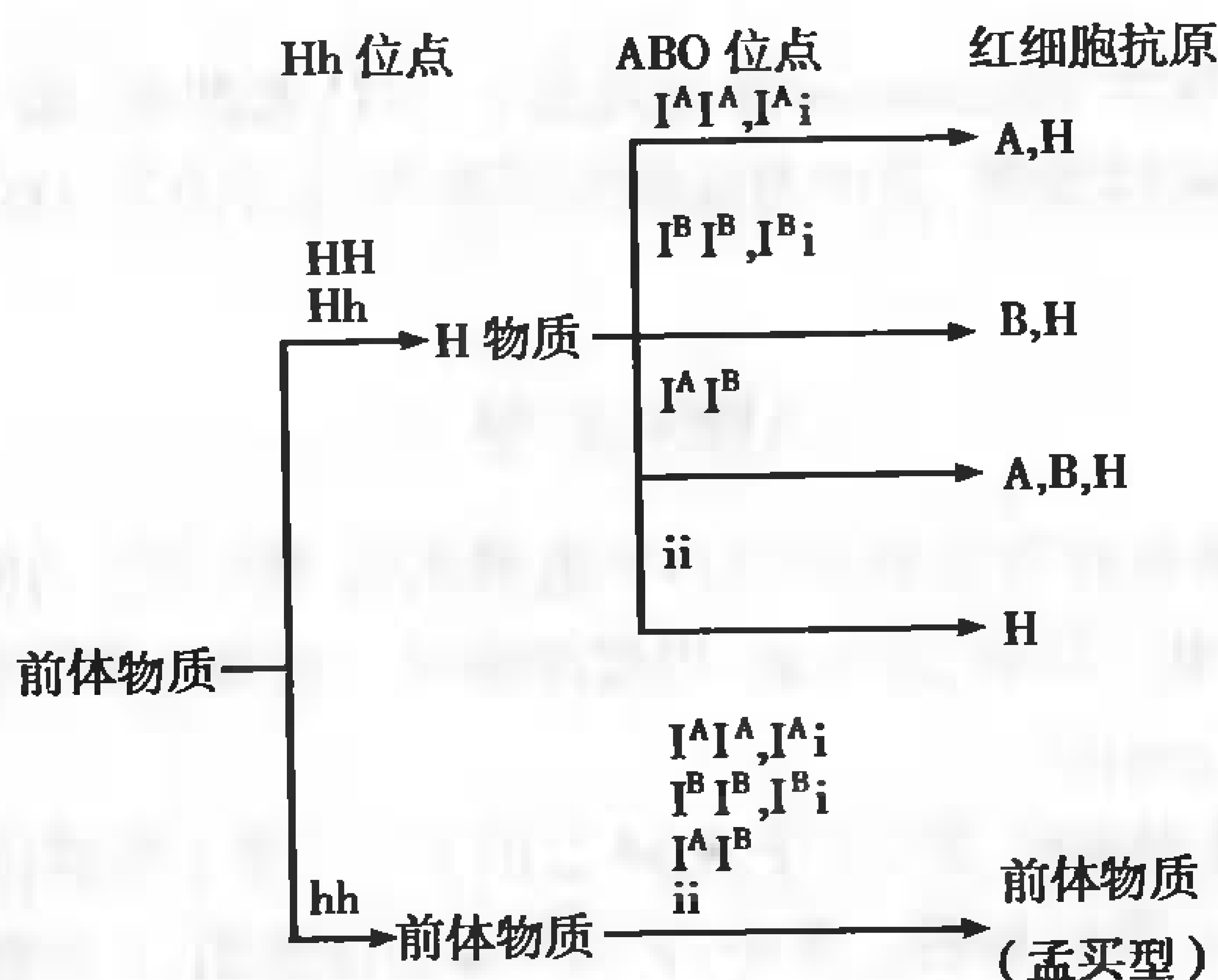


图 10－2　ABO 血型系统血型抗原的合成途径

能与 A、B 血型红细胞产生强烈凝集反应外，还可与 O 型红细胞发生凝集反应。当怀疑某人为 Oh 血型时，如其血清与 O 血型红细胞凝集，可以确定其为 Oh 血型。另外，植物凝集素(PHA) 可以与 O 血型红细胞发生强凝集，却不能与 Oh 血型红细胞凝集，故也可借此鉴别 Oh 血型与 O 血型。如用已知 Oh 血型血清鉴定 Oh 血型红细胞则更好。由于孟买型极其少见，所以当孟买型者需要输血时，选择同型的供血者是非常困难的。

人群中，80%的个体有 *Se* 基因，*Se* 基因的产物也是 *L*－岩藻糖转移酶，也能将 H 前体物质转化为 H 抗原，但其主要作用于分泌腺，使除脑脊液之外的各种体液中含有血型物质，称为分泌型；20%个体为 *sese* 基因型，不能编码 *L*－岩藻糖转移酶作用于分泌腺，所以，体液中不含有血型物质，故称之为非分泌型。

二、Rh 血型系统

Rh 血型系统是最复杂的一个血型系统，其重要性仅次于 ABO 血型系统。1940 年，Landsteiner 和 Wiener 用恒河猴（Rhesus monkey）的红细胞免疫家兔，所得抗血清能与约 85%白种人红细胞发生凝集反应，因此认为这些人红细胞含有与恒河猴红细胞相同的抗

原，故取名为 Rh 抗原，把相应抗体称为 Rh 抗体。

编码 Rh 抗原的基因定位在 1 号染色体(1p34.1 - p36)，*Rh* 基因由 2 个相关结构基因 *RHD* 和 *RHCE* 组成，彼此靠近，间隔仅 30 kb；2 个基因具有高度同源性，共长 69 kb，均含有 10 个外显子和 9 个内含子(图 10 - 3)。二者的基因产物均为含 417 个氨基酸的多肽，*RHD* 基因的缺乏或突变则不产生 D 抗原，*RHCE* 的基因产物经过不同的剪切可表达 CE、Ce、cE 和 ce 4 种抗原。由于 D 抗原的抗原性最强，故常以其存在与否确定为 Rh 阳性或阴性。与 ABO 血型不同，Rh 抗原没有天然抗体，机体只有接受 Rh 阳性的血液后，才能通过体液免疫产生抗 Rh 抗体。

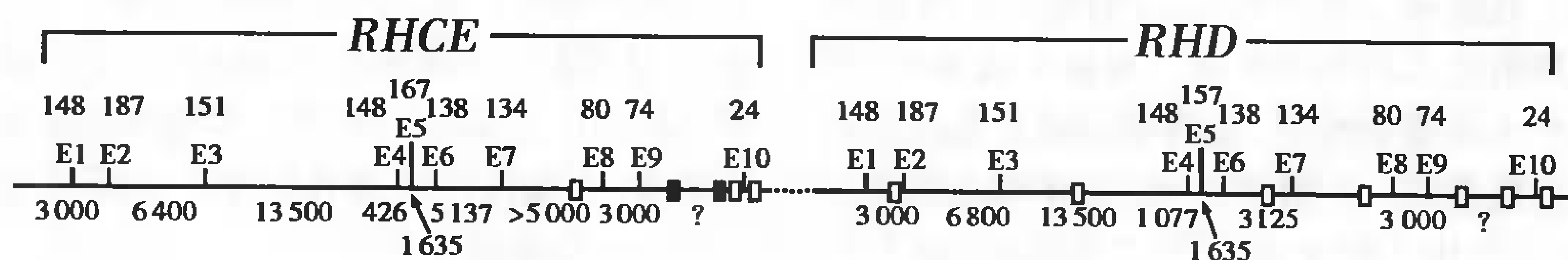


图 10 - 3 *RHCE* - *RHD* 基因结构
E：外显子；上方数字表示外显子大小，下方数字表示内含子大小

三、新生儿溶血病

新生儿溶血病(hemolytic disease of newborn, HDN)也称为胎儿有核细胞增多症(erythroblastosis fetalis)，发生在胎儿或新生儿时期。主要原因为母胎或母婴血型不合，母血中针对胎儿红细胞表面抗原的免疫抗体 IgG 通过胎盘进入胎儿血循环，发生同种免疫反应，导致胎儿或新生儿红细胞破坏而引起溶血，这是发生于胎儿或新生儿的、以溶血为主要损害的一种被动免疫性疾病。引起本病的血型抗体以抗 A 抗 B、抗 A、抗 B、抗 D 等为多见。抗 D 抗体引起者病情较重，Rh 血型系统中其他抗原的抗体引起者次之，抗 A、抗 B 抗体引起者病情较轻。国内以 ABO 血型不合引起者多见，Rh 血型不合引起者次之；白种人 Rh 溶血病发生率比黄种人高。

ABO 血型不合引起的新生儿溶血简称 ABO 溶血病，多发生于母亲为 O 型，婴儿 A 型或 B 型；如母亲为 AB 型或婴儿为 O 型则均不会发生溶血。由于 O 型妇女通常在孕前已接触过自然界中广泛存在的类似 A 型、B 型的血型物质，如某些植物、寄生虫及细菌等刺激，其血清中产生了相应的抗 A、抗 B 的 IgG 型抗体，孕期该抗体经胎盘进入胎儿血循环即会引起溶血，50% ABO 血型不合者第 1 胎即可发生溶血。

Rh 血型不合引起的新生儿溶血简称 Rh 溶血病，主要发生在 Rh 阴性母亲连续 2 次孕育 Rh 阳性胎儿的情况下。在第 1 次妊娠期，可有少量胎儿红细胞进入母亲血循环，但不足以引起母亲原发反应，初次致敏需要 0.5 ~ 1 ml 胎儿血液；即使能引起母亲致敏，引起原发反应，产生的抗体既少又弱，并且基本上是 IgM 型，不能通过胎盘；有超过 0.5 ml 胎血进入母体大多是由于分娩时的胎盘损伤、渗血所致，在这些情况下，虽然母亲可产生 IgG 型抗体，但由于胎儿已经娩出故不致引起溶血病。因此，Rh 溶血病一般不会在第 1 胎发生。当 Rh 阴性母亲再次妊娠 Rh 阳性胎儿时，由于在首次分娩时有超过 0.5 ml 胎血进入母体循环并已经产生过原发反应，故第 2 次致敏仅需 0.01 ~ 0.1 ml 胎血，并且在几天之内

就会产生大量 IgG 抗体，导致在胎儿期引起溶血。因此，Rh 溶血病症状随胎次增多而越来越严重。

极少数 Rh 溶血病发生在第 1 胎，其原因可能是孕妇孕前接受过 Rh 阳性的输血，也可能与孕妇是 Rh 阴性、而孕妇的母亲为 Rh 阳性有关。在后一种情况下，孕妇本人出生时已接受了 Rh 阳性母亲的抗原致敏。Rh 血型不合溶血病也可发生在母婴均为 Rh 阳性时，这是由抗 E、抗 C 或抗 e、抗 c 等引起，其中以抗 E 相对较多见。可通过给曾生过 Rh 阳性胎儿的 Rh 阴性孕妇注射抗 D 血清制剂的方法预防 Rh 溶血病。

新生儿溶血病的临床表现差别很大，重者可致胎死宫内，轻者仅在出生后皮肤轻微黄疸。其临床主要表现为：①贫血：Rh 溶血者，一般贫血出现早且重；ABO 溶血者贫血少，一般到新生儿后期才出现。重症贫血易发生贫血性心力衰竭。②高胆红素血症：由于溶血后产生大量胆红素，血清中未结合胆红素升高，婴儿皮肤严重黄染。因黄疸出现早且重易引起核黄疸。③肝脾肿大：由于贫血使器官组织缺氧，导致代偿性肝脾肿大。④胎儿水肿：患儿出生时全身水肿，皮肤苍白，常有胸、腹腔积液。⑤肌张力降低等。

第二节　人类主要组织相容性抗原

早在 20 世纪初，人们就已发现，不同种属或同种不同个体之间进行组织、器官移植时，会在移植后出现排斥反应致使移植失败。引起移植免疫反应的抗原称为移植抗原，又称组织相容性抗原。根据抗原性的强弱及引起移植排斥反应的强度，组织相容性抗原又可分为主要组织相容性抗原和次要组织相容性抗原两大类。编码主要组织相容性抗原的是一组连锁基因，称为主要组织相容性复合体（major histocompatibility complex，MHC），故主要组织相容性抗原也称为 MHC 抗原。MHC 抗原是引起移植排斥的主要抗原。各种动物都存在结构与功能与人类相似的 MHC 遗传区域，如小鼠、猪、狗、猴、黑猩猩的 MHC 分别是 H－2、SLA、DLA、RhLA、ChLA。人的 MHC 抗原几乎分布在人体所有有核细胞表面及血小板，但由于首先在人外周血白细胞上发现，并且白细胞上 MNC 抗原表达水平较高，目前又多采用外周血淋巴细胞来检测这类抗原的型别，故人类 MHC 抗原（分子）也称为人类白细胞抗原（human leukocyte antigen，HLA）；相应地，人类 MHC 称为 *HLA* 基因复合体，或 HLA 系统。虽然 HLA 是指抗原，但为了与 *HLA* 基因的提法相匹配，通常也将 HLA 称为 HLA 抗原（分子）。

一、*HLA* 基因复合体

HLA 基因复合体是迄今为止所知的人类最复杂的基因族。HLA 在调节机体免疫反应、破坏表达外来抗原的靶细胞方面有重要作用；通过 HLA 配型能提高移植物的存活率；作为一种遗传标记，*HLA* 基因用于有关疾病及人类遗传学的研究；在临床输血方面，HLA 的研究有助于提高成分输血的疗效及防止输血反应。总之，*HLA* 基因及其抗原的研究已广泛应用于基础医学、临床医学、预防医学、法医学等各个方面。

（一）*HLA* 基因复合体的组成与结构

HLA 基因复合体位于人类 6 号染色体短臂上（6p21.31），全长 3 600～4 000 kb。根据

HLA 基因编码的 HLA 抗原的分布、多态性及功能，将其分为 3 个区。*HLA* – Ⅰ类基因区位于 *HLA* 基因复合体的最远端，长约 2 000 kb；*HLA* – Ⅱ类基因区位于 *HLA* 基因复合体的近端，靠近着丝粒，长约 1 000 kb；*HLA* – Ⅲ类基因区位于Ⅰ类基因和Ⅱ类基因之间，长 600～1 000 kb(图 10 – 4)。

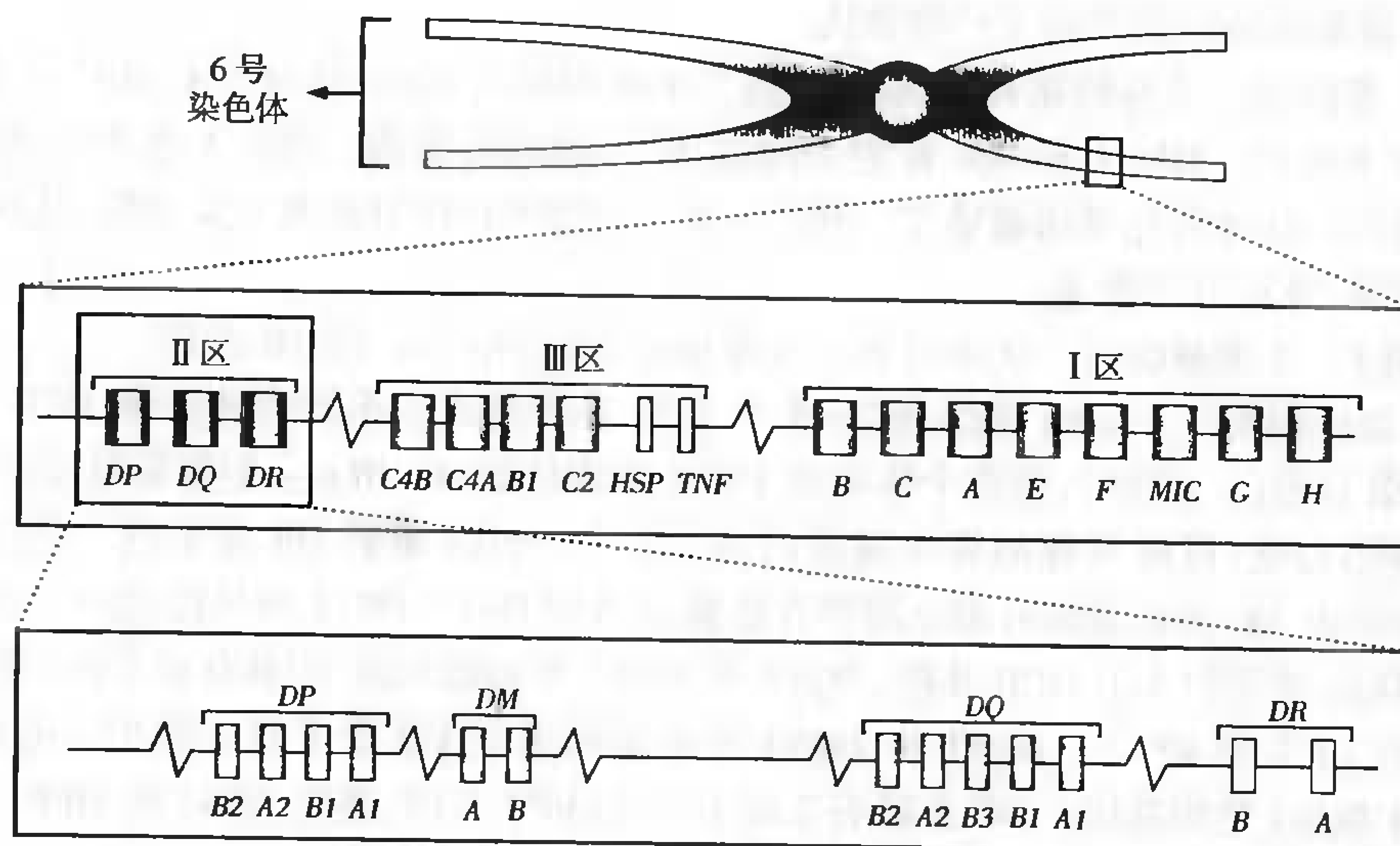


图 10 – 4 人类 *HLA* 基因复合体示意图

1. *HLA* – Ⅰ类基因　包括经典的 *HLA* – Ⅰ类基因(classical class Ⅰ gene, *HLA* – Ⅰ *a*)、非经典的 *HLA* – Ⅰ类基因(non – classical class Ⅰ gene, *HLA* – Ⅰ *b*)、已命名的假基因和 *MIC* 基因(MHC class Ⅰ chain – related, *MIC*)。

经典的 *HLA* – Ⅰ类基因是指最早发现的 *HLA* – *A*、*HLA* – *B* 和 *HLA* – *C* 基因，该组基因具有高度多态性，截至 2002 年 7 月，*HLA* – *A*、*HLA* – *B* 和 *HLA* – *C* 基因位点已命名的等位基因数分别为 248、486 和 113。*HLA* – Ⅰ类抗原分子是由非共价键连接的 2 条多肽组成的异二聚体膜糖蛋白，一条是 *HLA* – A 或 *HLA* – B 或 *HLA* – C 等位基因编码的 α 链(重链)，为 44 kDa 的跨膜糖蛋白，具有高度多态性；另一条是 15 号染色体的非 *MHC* 基因编码的 β 链(轻链)，为 12 kDa 的 β_2 微球蛋白(β_2m)，无多态性，游离于细胞外。HLA – Ⅰa 分子存在于几乎所有有核细胞表面，其生物学功能包括 3 个方面：一是作为组织排斥反应中的主要抗原，在移植免疫排斥中起主要作用；二是作为细胞毒性 T 淋巴细胞(cytotoxic T lymphocytes, CTL)用于识别抗原肽及杀伤靶细胞时受限制的抗原分子；三是呈递加工处理过的抗原肽。

非经典的 *HLA* – Ⅰ类基因包括 *HLA* – *E*、*HLA* – *F* 和 *HLA* – *G*，由于此类基因的多态性及编码产物分布的广泛程度均不及经典的 *HLA* – Ⅰ类基因，故称为非经典的 *HLA* – Ⅰ类基因。*HLA* – *E* 基因位点已命名的等位基因有 6 个，编码产物是自然杀伤细胞(natural killer cell)抑制性受体 CD94/NKG2 的特异性配体，在免疫调节中发挥重要作用。*HLA* – *G* 基因位点已命名的等位基因有 14 个，其编码产物仅表达在与母体组织直接接触的胎儿滋

养层细胞上，而这些细胞不表达经典的Ⅰ类和Ⅱ类抗原，其功能可能是作为一种配体分子，与NK等杀伤细胞抑制性受体结合，发挥抑制活性，保护胎儿滋养层细胞免受NK等杀伤细胞的作用。*HLA－F*基因的功能尚不清楚。

HLA－Ⅰ类基因中已命名的假基因有*HLA－L*、*HLA－H*、*HLA－J*、*HLA－K*和*HLA－X*。所有假基因均因突变而无产物表达。

*MIC*基因是一个基因家族，目前已发现5个基因位点，命名为*MIC A*、*MIC B*、*MIC C*、*MIC D*和*MIC E*。*MIC A*和*MIC B*为功能基因，其他为假基因。*MIC A*也具有高度多态性，目前已有51个等位基因被确定。MIC A分子主要表达在胃肠道上皮细胞，是NK细胞抑制性受体NKG2D的配体。

2.*HLA*－Ⅱ类基因区　又称D区。主要包括*DR*、*DQ*、*DP*和*DM*亚区。

*DR*亚区包括1个*DRA*基因和2～5个*DRB*基因位点。不同个体携带*DRB*基因位点数不同(*DRB*1～*DRB*9)，所有个体均有*DRB*1基因位点，是*HLA*－Ⅱ类基因区中最具多态性的基因位点，目前发现的等位基因已达271个。DRA编码DR分子的α链(重链)，*DRB*1编码β1链，两者组成可用血清学方法检测出的DR1～DR18特异性抗原。*DQ*亚区有2个*DQA*基因和3个*DQB*基因，*DQA*1和*DQB*1为功能基因，均具有多态性，其等位基因分别为20个和45个。*DQA*1和*DQB*1分别编码组成DQ分子的α链和β链；*DQA*2、*DQB*2和*DQB*3是假基因。*DP*亚区有2对*DPA*和*DPB*基因，其中*DPA*1和*DPB*1为功能基因，分别编码DP分子的α链和β链，*DPA*2和*DPB*2为假基因。

由于*DR*、*DQ*及*DP*基因所编码的抗原分布相似，并且均具有高度多态性，故被称为经典的*HLA*－Ⅱ类基因。DR、DQ、DP分子主要表达在胸腺上皮细胞、活化B细胞、巨噬细胞、树突状细胞活化等免疫细胞表面，其生物学功能包括两方面：一是作为移植排斥的主要抗原并参与免疫反应与免疫调节；二是呈递加工处理过的抗原肽给$CD4^+$的辅助性T细胞(T_H)。

*DM*亚区含有*DMA*和*DMB* 2个基因，也称为非经典的*HLA*－Ⅱ类基因，其编码的产物并不表达于细胞表面，但在细胞内*HLA*－Ⅱ类抗原分子装配过程中起重要作用，*DM*基因的缺失导致DR、DQ和DP分子不能表达。

3.*HLA*－Ⅲ类基因　是人类基因组中基因密度最大的区域，平均15 kb就有1个基因。此区的主要基因有：①补体基因(*C*2、*Bf*、*C4A*、*C4B*)；②21羟化酶基因(*CYP*21*A*、*CYP*21*B*)，其中*CYP*21*B*编码21羟化酶；③*HSP*基因，其产物为70 kDa的热休克蛋白；④肿瘤坏死因子(tumor necrosis factor，TNF)基因，其产物为细胞因子TNF－α和TNF－β；⑤淋巴毒素(lymphotoxin，LT)*A*基因和*B*基因，分别编码淋巴毒素α和淋巴毒素β。*HLA*－Ⅲ类基因的编码产物主要存在于血清和体液中。

(二)*HLA*基因复合体的遗传特性及多态性

1.单元型遗传　*HLA*基因复合体是一组紧密连锁的基因群，其基因在同一染色体上的组合就构成了单元型(haplotype)。由于1条染色体上*HLA*基因复合体各位点的距离非常近，很少发生同源染色体之间的交换，故*HLA*单元型作为一个单位遗传给子代，称单元型遗传。子代的*HLA*单元型一个来自父亲，一个来自母亲。假如父亲2条染色体上的*HLA*单元型为*a*和*b*，母亲的*HLA*单元型为*c*和*d*，则子代可能有4种*HLA*单元型组合

(*HLA* 基因型):*ac*,*ad*,*bc*,*bd*。同胞之间 *HLA* 基因型完全相同或完全不同的可能性都是 1/4;1 个单元型相同的可能性是 1/2。而子代和亲代总有 1 个相同的单元型。

2.连锁不平衡 理论上,在随机组合的情况下,一个 *HLA* 基因复合体各位点的等位基因与另一个或几个位点的等位基因在某一单元型出现的频率应等于各位点基因频率的乘积。然而,在很多情况下,预期的单元型频率往往与实际检测的频率相差很大,在同一条染色体上 2 个或多个等位基因同时出现的频率高于或低于其随机组合频率,这种现象称为连锁不平衡(linkage disequilibrium)。例如,白种人常见 *A*1 - *B*8 - *DR*3 - *DQ*2 单元型,*A*1 的基因频率为0.12,*B*8 的基因频率为 0.17,*A*1 - *B*8 单元型的预期频率为 0.12×0.17 =0.02,但实际为 0.09。另外,北欧人 *HLA* - *A*1、*B*8,*HLA* - *A*8、*B*7 2 个单元型最常见,黄种人则 *HLA* - *A*9、*B*15 单元型较常见,我国汉族人 *HLA* - *A*2、*B*46,*HLA* - *A*11、*B*40 单元型最常见。*HLA* 基因连锁不平衡的发生机制目前尚不清楚,但已经发现某些疾病的发生与 *HLA* 基因中某些特定的等位基因密切相关;某些连锁不平衡倾向于出现在某些地区、某些人种或某些民族。深入探讨连锁不平衡的发生机制无疑将有助于对某些疾病的诊断和治疗。

3.多态性 *HLA* 基因复合体是人体最复杂的基因系统,呈高度的多态性,主要原因是 *HLA* 基因复合体的多数基因位点存在复等位基因。普通多态性基因的复等位基因不超过 10 个,而 *HLA* 基因的多态性要复杂得多,如 *HLA* - *A*、*HLA* - *B* 及 *HLA* - *C* 的复等位基因分别为 248、486 和 113 个,*HLAB*1 位点复等位基因为 271 个,而且这些基因位点还不断有新的复等位基因被发现。其次,各个基因位点的基因组合是随机的,人群中大多数个体 *HLA* 基因位点是杂合子,其不同的等位基因编码的产物可表达在同一细胞膜上,即共显性遗传。因此,*HLA* 表型呈现出多样化。*HLA* 基因的多态性使得机体能对各种病原体产生适宜的免疫反应;同时,这种高度多态性又是一个很好的人类遗传学标记,使每个个体都带有自己独特的一套生物学身份证,对人类学研究具有重要意义。但另一方面,在进行器官移植时,由于 *HLA* 的多态性使得组织配型完全相合的概率大大减少,给器官移植造成困难。

(三)*HLA* 等位基因/抗原的检测与命名

1.HLA 抗原特异性检测 *HLA* - *A*、*B*、*C*、*DR*、*DQ* 和 *DP* 位点基因表达的抗原可用血清学方法检测出宽特异性。书写时,在基因位点后直接写上特异性(抗原)即可,如 HLA - A2、HLA - DQ3、HLA - DR9 等。用细胞学方法检测出的抗原特异性,书写时均需加上"w",C 位点抗原后也均需加上"w"以区别补体成分,如 HLA - DPw、HLA - Qw 等。

2.*HLA* 等位基因的检测与命名 分子生物学方法目前已被应用于组织配型,在 PCR 基础上发展了各种 DNA 分型技术。应用这些技术进行 *HLA* 等位基因分型后,发现同一个血清学特异性可被数个甚至数十个等位基因所编码,如编码 HLA - A2 抗原的等位基因至少有 39 个,编码 HLA - DR4 的等位基因至少有 36 个。这说明血清学方法检测的 HLA 抗原的特异性不强,即宽特异性。鉴于此,1989 年,世界卫生组织命名委员会制定了 *HLA* 基因的命名原则:对一个等位基因命名,首先用大写斜体字母表示基因位点,然后接"*",再用 4 位数字代表该等位基因名。例如:*HLA* - *DRB*1 * 0405、*HLA* - *B* * 2701、*HLA* - *DQA*1 * 0301 等,*DRB*1、*B*、*DQA*1 均指 *HLA* 基因位点,"*"后前 2 位数表示该基因位点的

复等位基因,应尽可能与该等位基因的血清学特异性相同,后2位数字则代表该等位基因的亚型。如一等位基因在命名后,又在DNA序列水平发现与之相似、仅有微小差异的等位基因时,则在原等位基因亚型后再加上第5个数字以区别,如 *HLA - B* * 51011 及 *HLA - B* * 51012。

二、HLA与器官移植

器官移植(transplantation of organs)是重要的临床治疗手段,自1954年Murray第1次施行同卵双生姐妹之间的肾移植获得长期存活以来,随着对移植免疫与免疫遗传认知程度的提高,免疫遗传学理论和技术的发展及强有力的免疫抑制药物的应用,使得器官移植、骨髓移植的成活率大大提高。

(一)器官移植与免疫排斥

每个个体都有自己的免疫特异性,在全身的组织、细胞、器官上均有特异的、通过遗传而获得的特异性抗原。当带有不同特异性抗原的异己组织、细胞、器官移植进入机体后,就会刺激机体的免疫系统,产生针对移植物的免疫排斥反应。供者与受者的 *HLA* 基因的相似程度直接反映了两者的相容性,供者-受者间的HLA相似性越高,移植成功的可能性越大。自体器官移植或同卵双胎之间进行移植几乎不发生排斥反应;同胞或亲子之间因有1条 *HLA* 单元型相同,移植成功的可能性也较大;无任何亲源关系的个体间器官移植的成功率要低得多。

(二)组织配型

临床器官移植的供者和受者间多为无亲缘关系,为了降低移植排斥反应,延长移植物的存活时间,移植前必须进行组织配型,了解供受双方组织相容性程度,尽可能为受者寻找组织相容性最合适的个体。由于ABO血型抗原和HLA抗原是最重要的组织相容性抗原。因此,组织配型的原则是使供者和受者的ABO血型抗原和HLA抗原相同,或者供者没有受者所缺的抗原。

1.ABO血型抗原配型　ABO血型抗原不仅存在于红细胞膜上,在其他组织细胞膜上也有表达。因此,ABO血型抗原相容是器官移植成功的先决条件。器官移植的ABO血型抗原配型的原理、方法与临床输血配型相同,即O血型受者只接受O血型移植物;A血型受者接受A血型或O血型移植物;B血型受者接受B血型或O血型移植物;AB血型受者可接受各血型移植物。一般情况下,不管是活体移植物还是尸体移植物,血型不相容就不应该进行移植。

2.HLA配型　临床器官移植的供者与受者 *HLA* 基因位点完全相同的可能性几乎不存在。供者和受者间多无血缘关系,在这种情况下,*HLA* 基因相同的位点越多,相容性越好,移植物长期存活的可能性就越大。只要存在1~2个 *HLA* 基因位点相同,就比 *HLA* 基因位点完全不相同者之间的移植成功率大得多。由于 *HLA* 基因呈单元型遗传,因此,在近亲中寻找HLA相同或半相同的概率远远大于无亲缘关系的供体。如果有可能,供者最好是父母、子女或同胞中的健康志愿者,这样可以在移植前对双方进行较全面的检测和交叉配型试验,尽可能选择有较多的 *HLA* 基因位点相同者作为供者。在近亲中,除同卵双生者外,供体首选HLA相同的同胞,次选HLA半相同的同胞或双亲或子女。*HLA* - Ⅱ类

基因的相合对于避免和减轻移植物抗宿主反应(GVHR)和肾等移植器官的长期存活非常重要,这是因为供者和受者 HLA-Ⅱ类抗原相合或相合较好,受者的 TH 细胞不发生或只发生轻度活化和增殖,产生细胞因子的量不足于诱导 B 细胞和 CTL 的增殖和分化。因此,在 HLA-Ⅱ类抗原相合前提下,有时即使Ⅰ类抗原不相合,受者也不会发生明显的针对移植抗原的免疫反应和 Tc 的杀伤作用。

在器官移植方面,肾移植开展较早,临床意义也较大。供者与受者间组织相容性程度与肾移植物的存活率相关。据报道,肾移植 3 年、5 年和 10 年存活率,在供受双方 *HLA* 完全相同(同卵双生)时分别为 80%、81%和 70%;1 个单元型一致(亲子间)时分别为 65%、55%和 40%,在完全不同时分别为 50%、36%和 20%。另有资料表明,供受双方 *HLA-A*、*B* 和 *DR* 位点抗原全部相符时,1 年移植肾存活率高达 93%;在 *HLA-DR* 相符,*HLA-A*、*B* 2 个位点中 1 个不符时,1 年移植肾存活率仍高达 89%;相反,*HLA-A*、*B* 位点完全相符而 *HLA-DR* 位点中 1 个不相符时,移植肾 1 年存活率下降到 70%。这说明 *HLA-DR* 位点的组织相容性对移植肾成活更重要。目前,倾向于组织配型时 *HLA-DR* 位点和 *HLA-B* 位点抗原同时检测。

骨髓移植对 *HLA* 组织配型要求更加严格,因为骨髓中含有免疫活性细胞,可以识别宿主细胞并产生免疫反应。如 *HLA* 不符,尤其是 *D* 位点不同会引起严重的移植物抗宿主反应(GVHR)。由于无血缘关系个体的 *HLA* 的高度多态性,无关个体间 *HLA* 基因型完全相同的概率极低,骨髓库的建立将为骨髓移植在临床上更好地应用提供保障。

免疫抑制剂(如环孢素 A)在器官移植中的应用,使得排斥反应缓解,人们曾一度认为选择 HLA 相容性供者已可有可无。然而,近年来组织配型的重要性再次使人们取得共识,特别认识到 *HLA*-Ⅱ类基因,尤其是 *HLA-DP* 的相容性对移植器官存活仍是极为重要的,对骨髓移植避免 GVHR 尤为如此。*HLA* 多态性使Ⅰ类和Ⅱ类基因位点完全相合的可能性几乎仅见于同胞之间,供者的来源相当受限。目前的解决办法有 2 个:一是采用单元型相同的家庭成员为供者;二是选择 HLA 表型相同的无关供者(unrelated donor,URD)。近来年,世界范围内 URD 库急剧扩大,已有 50 万以上 URD 的 HLA-*A* 和 -*B* 的分型记录在案,其中的 18 000 人还有 -*DR* 和 -*DQ* 的分型资料。在由 URD 提供器官的器官移植中,组织配型的重要性更显突出。我国推行独生子女政策,对 URD 库的建立似更为迫切。

三、HLA 与疾病的相关性

(一)HLA 与疾病的关联

在 1967 年第 3 届国际组织相容性专题讨论会上,Amiel 首次报道霍奇金病患者中 HLA 抗原 4C 频率高于正常对照组,提示该病与 HLA 相关联。该发现引起国内外学者的广泛关注。自那时起,世界各国掀起了寻找 HLA 与疾病关联的热潮。迄今为止,已对 500 多种疾病进行了分析,并发现 60 余种疾病与 *HLA* 基因相关联。

近年来,HLA 与疾病关联的研究不再局限于个别的 *HLA* 基因或其产物,而是扩展到整个 *HLA* 基因复合体的某一段,称其为 MHC 扩展的单元型(*MHC*-extended haplotype,*MHC-EH*)。研究 *MHC-EH* 与疾病的关联比研究单个基因或单个表型更有意义。例如,北美白种人中 HLA-DR4 和 HLA-B62 抗原携带者患类风湿性关节炎的相对危险率

(*RR*)分别为5.0和2.5,而*HLA - EH*(*B62 - Bf - C4B - DR4*)携带者的*RR*为16.1,高于*B62*与*DR4*二者*RR*之和。患者中*HLA - EH*(*B62 - Bf - C4B - DR4*)的频率为9%,而对照组仅为0.6%。这说明,某些疾病的易感基因可能不是单一基因位点的等位基因,而可能是多个基因位点等位基因组成的基因群。以下简介部分与HLA有关联的疾病的研究进展。

1.强直性脊椎炎(ankylosing spondylitis,AS)　AS是第1个被证实与HLA抗原强相关的疾病。研究证实,AS发病率与HLA - B27在群体中的频率有关,HLA - B27的抗原携带者罹患强直性脊椎炎的风险较不携带HLA - B27抗原者高出90倍;在同卵双生子中患病一致率研究强烈提示,AS是一种基因遗传病,遗传率约为98%。进一步的研究发现,与AS相关联的8个*HLA - B27*等位基因(*HLA - B* * 2701 ~ *HLA - B* * 2708)中,*HLA - B* * 2705最常见,而*HLA - B* * 2703则从未发现,提示*HLA - B* * 2703基因可能是AS的保护基因。

2.胰岛素依赖性糖尿病(IDDM)　IDDM的发病有明显的种族差异,群体发病率为0.4%。当同胞之一患病时,其他同胞的发病风险增加15倍,发病率增至6%;同胞*HLA*基因型完全相同时,其发病的一致率高达15% ~ 25%。这些都说明IDDM是一种多基因遗传病。研究表明,IDDM的易感基因至少有15个,其中,*IDDM1*是主基因。这里所说的主基因并非指单一的基因位点,而是指包括*DR*、*DQ*在内的*HLA*基因的一组连锁位点,主要有*DRB1*、*DQA1*和*DQB1*。这些位点的基因有的具易感效应,有的具保护效应,并且显示出等级之差。这种效应以*DRB1* * 04等位基因最为突出,即易感性自强到弱依次为:0405、0402、0401;保护作用自小到大依次为:0403、0406、0408。另外,*DR*、*DQ*编码的β肽链上第57位氨基酸为天冬氨酸时也具有保护作用。事实上,HLA关联的IDDM的发病风险是由DR和DQ分子之间复杂的相互作用决定的。

3.类风湿性关节炎(rheumatoid arthriris,RA)　1977年,McMichael等首先报道RA的发病与*DR4*相关。进一步的研究表明,不同群体的RA相关位点不同。例如,在*DR4*频率较高的白种人中,RA患者的*DRB1* * 0401、* 0404、* 0408基因频率增高;而在频率较低的群体中,RA患者*DRB1* * 0101基因频率增高;日本人中,RA患者的*DRB1* * 0405基因频率增高;而在所有群体中,*DRB1* * 040均与RA发病无关。通过对这些基因所编码的HLA抗原β链氨基酸序列分析,发现它们都有共同的结构特点,即β链第3高变区70到74位氨基酸为QKRAA或QRRAA。从HLA - DR分子的主体结构看,该区氨基酸恰好构成由β链α螺旋形成的"分子袋"侧壁的一部分,该"分子袋"能有效地与某些多肽结合,而这些多肽则与其他HLA抗原不能结合或不能很紧密地结合。这部分与RA关联的分子本身可能代表了RA发病的遗传学基础。

4.系统性红斑狼疮(systemic lupus erythematosus,SLE)　SLE是一种多基因遗传病,与HLA - Ⅲ类抗原即补体成分的缺失有明显关联。正常人群C2缺失率为1% ~ 2%,SLE患者缺失率则达6%;1个C4A缺失时,该病的*RR*为3,2个C4A缺失时,*RR*增至17。DR2与C4A缺失同时存在时,则*RR*增至25。

(二)HLA与疾病关联的机制

HLA与某些疾病相关的机制目前尚不完全清楚,可能与以下几种假说有关。

1.分子模拟假说　某些病原微生物的抗原与 HLA 抗原分子结构相似，即为共同抗原，以致机体对这种病原微生物产生交叉免疫耐受，不能产生有效的免疫应答，使得病原微生物不能有效地被清除而致病，或者病原微生物刺激机体产生相应的抗体，损伤了具有共同抗原的组织细胞。例如，强直性脊椎炎和 Reiter 病患者细胞表面 B27 抗原与肺炎杆菌蛋白成分有一段共同的氨基酸序列。

2.受体学说　HLA 抗原可能作为某种病原体的受体，二者结合可造成机体组织的损害。例如，小鼠 MHC Ⅱ类抗原是乳酸脱氢酶病毒的受体。

3.免疫功能失调假说　*HLA* - Ⅱ类基因中包含有控制免疫应答能力和调节免疫功能的基因，其产物影响巨噬细胞提呈抗原或与其他细胞间的相互作用。特定的Ⅱ类基因型别的异常产物影响机体对免疫功能的调节，从而产生异常的免疫反应，使机体对某些疾病易感。

4.连锁不平衡学说　*HLA* 基因与真正的疾病易感基因连锁不平衡。*HLA* 型别仅是一种可供检出的遗传标志，而与 *HLA* 基因位于同一条染色体上的真正的疾病易感基因尚未被发现。

第三节　抗体遗传

免疫球蛋白(immunoglobulin，Ig)是 B 淋巴细胞识别抗原后，增殖、分化的浆细胞产生的能与相应抗原特异性结合的一种蛋白质，分为分泌型和膜型。分泌型又称为抗体，主要存在血液中，但也见于其他体液和外分泌液；膜型存在于 B 淋巴细胞膜上，称为 B 细胞抗原受体(BCR)。二者的分子结构及遗传控制相同，因此，免疫球蛋白通常称为抗体。抗体具有高度特异性，因而表现出高度多样性。

一、免疫球蛋白的基本结构

人类 Ig 分子基本结构由 2 条相同的轻链和 2 条相同的重链组成，重链与轻链间通过二硫键连接形成 Ig 分子单体。根据重链恒定区分子结构和抗原特异性，人类 Ig 分为 5 类：即：IgG、IgA、IgM、IgD 和 IgE，其中相应的重链(heavy chain，H 链)为 μ 链、γ 链、α 链、δ 链和 ε 链。轻链(light chain，L 链)分为 κ 链与 λ 链两型。一种哺乳动物只有 1 种 κ 链，但 λ 链分为 4 个亚型($\lambda1$、$\lambda2$、$\lambda3$、$\lambda4$)。一种特异性的 B 淋巴细胞只能合成 1 种 L 链，要么是 κ 链，要么是 λ 链中的 1 种，也只能合成其中的 1 种。一个天然 Ig 分子上的 2 条轻链的型别总是相同的。5 类 Ig 中，每类 Ig 都可以有 κ 链或 λ 链。其中，IgG、IgD 和 IgE 是单体，IgA 是二聚体，IgM 是五聚体。

根据 L 链和 H 链的结构与生物学功能，将其分为可变区和恒定区。L 链 N 端 1/2 (V_L)、H 链 N 端 1/4(V_H)区域内的氨基酸排列顺序随抗体的特异性不同而变化，称为可变区(variable region，V 区)。抗体的多样性就是可变区决定的，可变区也是抗原结合的部位。L 链 C 端 1/2(C_L)和 H 链 C 端 3/4(C_H)区域，在同一种属动物中，氨基酸种类、数量和排序是比较恒定的，故称为恒定区(constant region，C 区)。C_H 又包括 C_H1、C_H2、C_H3 3 个功能区，有的 C_H(μ 链、ε 链)还有 C_H4 功能区。

二、免疫球蛋白基因结构及重排

Ig分子由3个不连锁的κ链基因、λ链基因和H链基因编码。κ链基因位于2p11,λ链基因位于22q11,H链基因位于14q32.3。每个Ig基因群长80万至200万bp,是普通基因长度的50倍以上。

(一)H链基因的结构和重排

胚胎细胞中H链基因总长约130 bp,由L_H-V_H、D_H、J_H、C_H 4种基因片段组成。第1部分含L_H-V_H基因片段,约65个,长约2 000 kb,其中L_H是先导肽序列基因片段;第2部分为D_H基因片段,约30个,长约80 kb,位于V_H与J_H之间;第3部分是6个J_H基因片段,长8 kb;第4部分是C_H基因片段,依次排列有9个功能性基因:$C\mu - C\delta - C\gamma3 - C\gamma1 - C\alpha1 - C\gamma2 - C\gamma4 - C\varepsilon - C\alpha2$。

在胚胎发育过程中,B淋巴细胞逐步分化、发育,在特异性重组酶的作用下,被分隔的无功能的基因片段经重排连接成一个完整的有转录功能的活性基因。重排时,首先发生的是D_H与J_H的连接,形成D-J基因片段,然后再与V_H连接,形成V-D-J基因片段,最后与C基因片段连接,形成完整的H链基因(图10-5)。

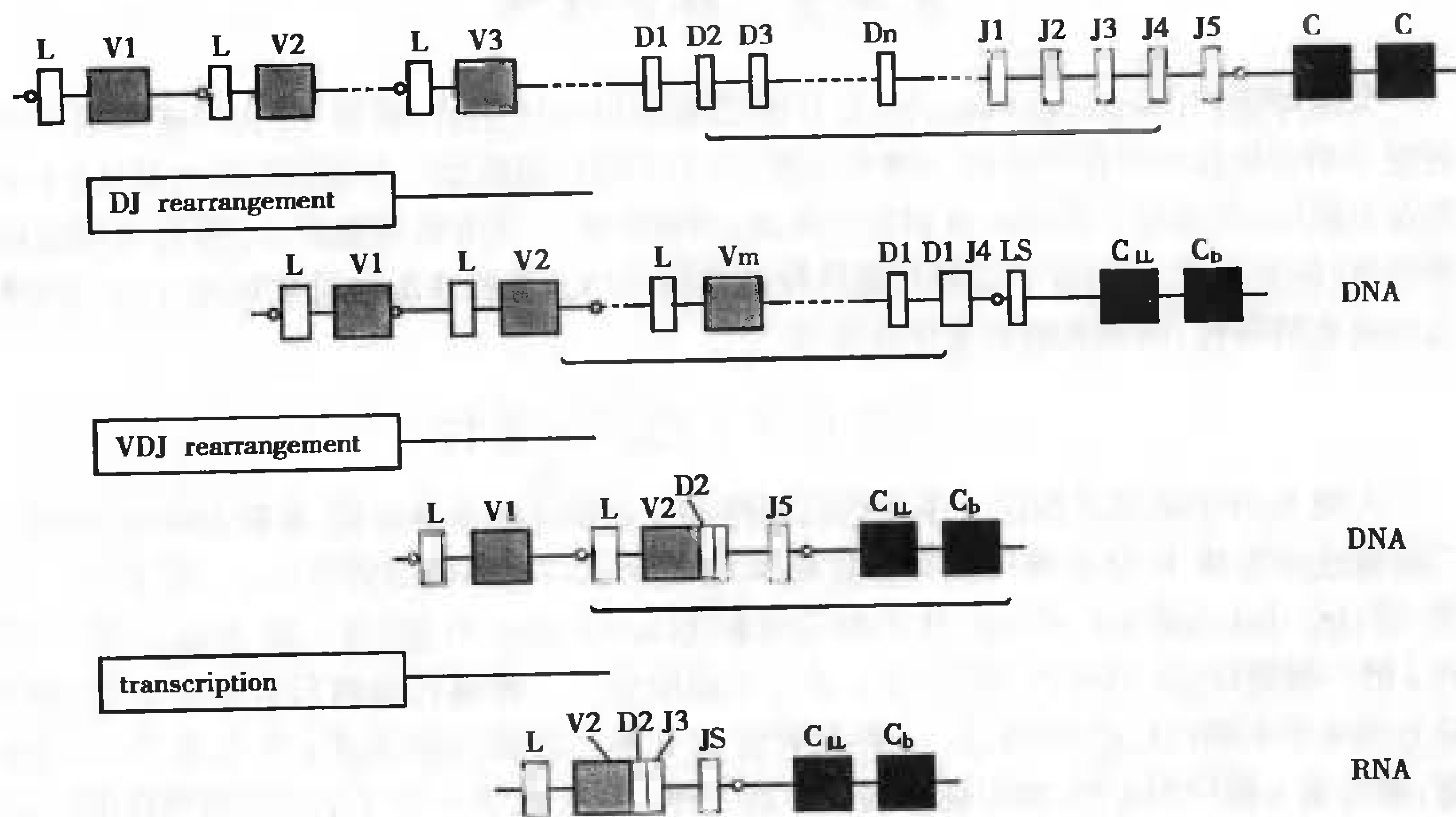


图10-5　重链基因重排过程

(二)L链基因的结构和重排

人胚胎细胞κ链基因和λ链基因均分为L-V(Lκ、Lλ,Vκ、Vλ)、J(Jκ、Jλ)、C(Cκ、Cλ)3种基因片段,无D基因片段。H链基因重排可诱导L链基因片段重排,首先是κ链基因重排,如果κ链基因重排无效,随即发生λ链基因重排。

人胚胎细胞κ链基因全长约200万bp,Lκ-Vκ基因片段约50个,Jκ基因片段5个,Cκ基因片段只有1个。重排时,在特异性重组酶的作用下,首先是Lκ-Vκ与Jκ拼接,然后,

L–V–J 基因片段再与下游的 Cκ 连接，构成有转录活性的 κ 链基因，其转录产物经过加工，成为成熟的 mRNA，从而翻译出特异的 κ 链。λ 链基因的重排与 κ 链基因重排相似（图 10–6）。

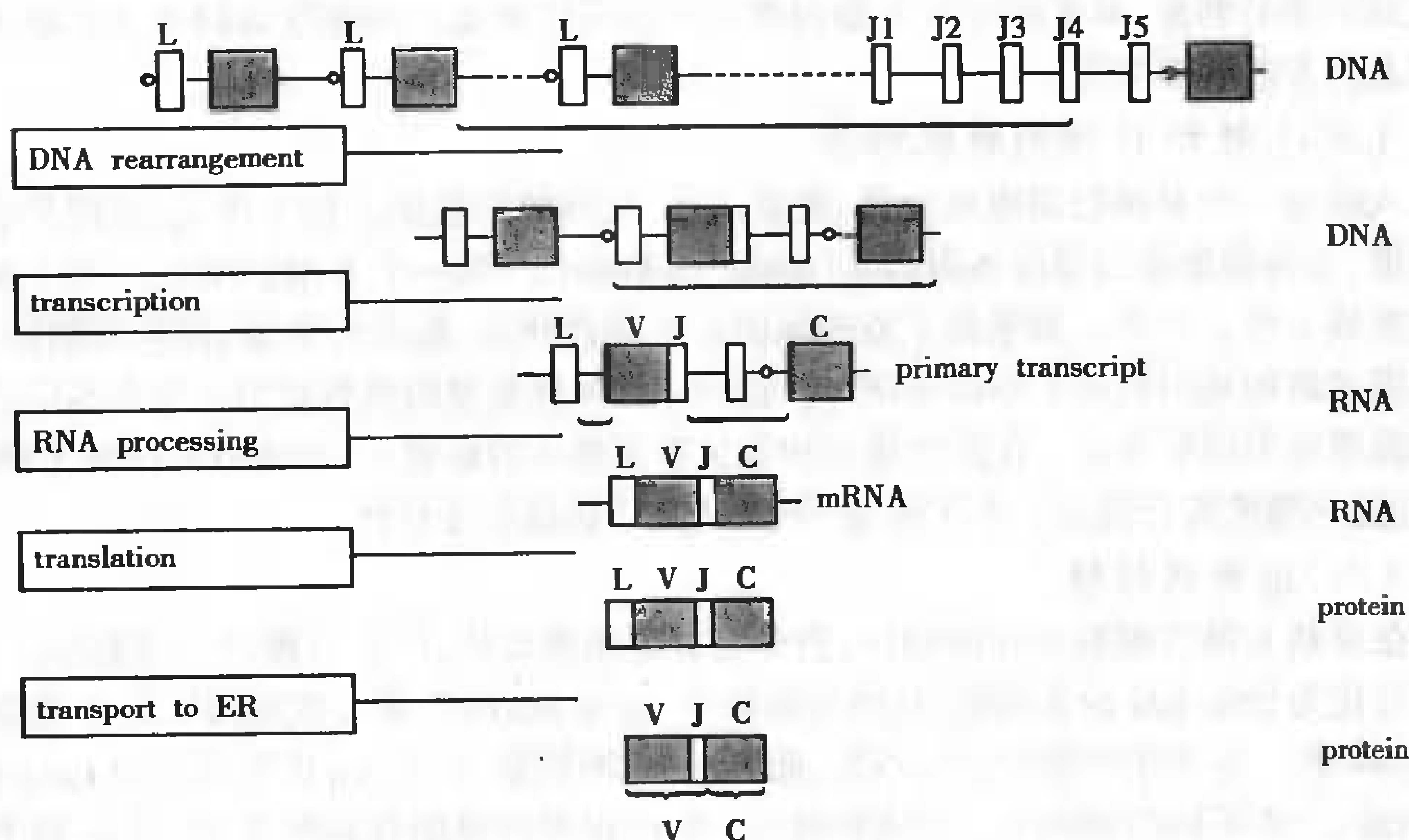


图 10–6 κ 轻链基因重排过程

三、免疫球蛋白的多样性

机体受外界环境中各种抗原刺激可产生相应的特异性抗体。外界环境中的抗原种类极多，每种抗原上又有不同的抗原决定簇。因此，受抗原刺激产生的特异性抗体的数量必然相当大，据估计，人体内抗体的种类可达 10^9 ~ 10^{11}。人类基因组仅含 3 万 ~ 5 万个基因，不足以表达如此巨大数量的不同特异性的抗体，抗体多样性的产生有以下几方面的机制。

（一）胚系基因片段的随机组合

胚胎发育过程中，前 B 淋巴细胞的 H 链基因、κ 链基因、λ 链基因都由众多的基因片段组成，当接受外界抗原的刺激后，这些基因发生重排。每一个 Ig 的 H 链基因或 L 链基因均由 V、(D)、J、C 基因片段中各选择 1 个组合而成，由于 V、D、J 和 C 基因片段的多数量性和选择的随机性，故重排可产生大量的序列不同的基因，从而表达大量的特异性免疫球蛋白。

（二）V–(D)–J 连接的多样性

在基因重排过程中，κ 链基因和 λ 链基因的 V–J 连接点，H 链基因的 V–D–J 连接点常发生核苷酸丢失现象，故可产生不同的连接方式，增加了重排基因的多样性。

（三）N 区的插入

在 H 链基因片段重排过程中，在 D_H 的两侧即 $V_H - D_H$ 或 $D_H - J_H$ 连接处额外插入几个核苷酸，构成 N 区。由于额外插入了 N 区（N region），导致移码突变，使插入部位以及下

游的密码子改变，引起编码的氨基酸改变，从而增加了抗体的多样性。

（四）体细胞高频突变

体细胞高频突变是指 B 淋巴细胞发育后期，次级淋巴器官生发中心的成熟 B 淋巴细胞受到抗原的刺激，其重排后的 V 基因所发生的基因突变。体细胞高频突变扩展了原有胚系基因片段的多样性。

（五）L 链和 H 链的随机组合

人体每一个 B 淋巴细胞在分化、发育过程中，同源染色体上的 2 个 *Ig* 基因只有 1 个有功能，这种现象称为等位基因排斥（allelic exclusion）。同一个 B 淋巴细胞内有 2 类轻链基因家族 κ 和 λ，只有 1 种基因家族的基因处于活性状态，能进行重组、转录和翻译，另一种基因家族的基因则处于非活性状态。这种一种家族的基因活性对另一家族基因活性的抑制现象称为同型排斥。在同型排斥和等位基因排斥的影响下，表达的 H 链和 L 链（κ 链和 λ 链）的随机组合构成了不同的 Ig 分子，增加了抗体的多样性。

（六）Ig 类别转换

在成熟 B 淋巴细胞的存活期内，当受到抗原刺激活化，产生 B 淋巴母细胞后，一部分细胞分化为分泌 IgM 的浆细胞，其他细胞则对 *Ig* 基因进行“第 2 次重排”，这种现象称为 Ig 类别转换。Ig 类别转换时，V_H 不变，而 C_H 发生不同重排，其 V_H 可和不同的 C_H（Cγ、Cα、Cε）相连，产生不同类别的 Ig。类别转换后，该 Ig 识别抗原的特异性未变，而 Ig 的类或亚类发生改变。Ig 可能是通过缺失模式和 RNA 剪接 2 种机制来实现类别的转换，类别转换的实质就是在每个 C_H 外显子前面特异部位剪接，使不同的 C_H 外显子与 V－D－J 组合。

小　结

ABO 血型抗原由位于 9q34 上的 I^A、I^B 和 *i* 复等位基因和位于 19 号染色体的 *H*、*h* 及 *Se*、*se* 等位基因通过不同组合以共显性方式编码而成，其中 I^A、I^B、*H*、*Se* 为显性基因，*i* 为隐性基因，*h* 和 *se* 为无效基因基因型。*hh* 表现为孟买型，*Se* 为分泌型，*se* 为非分泌型。Rh 血型抗原由位于 1p34.1－p36 上的 2 个相关结构基因 *RHD* 和 *RHCE* 所编码。*RHD* 基因表达 D 抗原，*RHD* 基因的缺乏或突变则不产生 D 抗原，以 D 抗原存在与否确定 Rh 阳性或阴性。*RHCE* 的基因产物经过不同的剪切可表达 CE、Ce、cE 和 ce 4 种抗原。当母胎/母婴血型不合时，可引起新生儿溶血病，以 ABO 血型不合者多见，约 50% 发生在第 1 胎；Rh 血型不合者较少见，多发生在第 2 胎，症状较严重。

人类白细胞抗原（HLA）是由 6p21.31 上的一组连锁基因编码。*HLA*－Ⅰ类基因位于 *HLA* 基因复合体最远端，具有高度多态性，编码产物是组织排斥中的主要抗原、杀伤性 T 淋巴细胞用于识别抗原肽及杀伤性淋巴细胞受限制的抗原分子、自然杀伤细胞的配体分子，并有呈递加工处理过的抗原肽的功能。*HLA*－Ⅱ类基因位于复合体的近端，编码产物是移植排斥的主要抗原，参与免疫反应与调节，并有呈递加工处理过的抗原肽给 $CD4^+$ 辅助性 T 细胞的功能。*HLA*－Ⅲ类基因位于Ⅰ类基因和Ⅱ类基因之间，编码一些补体及细胞因子。*HLA* 基因复合体具有单元型、连锁不平衡、多态性的遗传特性。器官移植时，通过 ABO 血型抗原配型和 HLA 配型，可降低免疫排斥反应，提高移植物存活率。HLA 与强

直性脊椎炎、胰岛素依赖性糖尿病、类风湿性关节炎及系统性红斑狼疮等疾病的发病相关。

抗体是B淋巴细胞识别抗原后，增殖、分化的浆细胞所产生的能与相应抗原特异性结合的免疫球蛋白(Ig)，由位于14q32.3的H链基因和位于2p11的κ链基因或位于22q11的λ链基因编码。H链基因由L_H-V_H、D_H、J_H、C_H 4种基因片段组成，κ链基因和λ链基因均由L-V(Lκ、Lλ，Vκ、Vλ)、J(Jκ、Jλ)、C(Cκ、Cλ)3种基因片段组成，胚胎发育过程中进行基因重排。H链基因重排后诱导L链基因重排，首先是κ链基因重排，如果无效则发生λ链基因重排。由于胚系基因的随机组合、V-(D)-J连接的多样性、N区的插入、体细胞的高频突变、重链轻链的随机组合以及Ig的类别转换，使得抗体的表达具有多样性，满足机体的免疫需要。

(施旭东　孙贝加　蔡绍京)

第十一章　肿瘤遗传学

肿瘤遗传学是研究肿瘤发生与遗传因素之间关系的医学遗传学分支学科。肿瘤细胞是积累了不同基因突变的体细胞，突变导致细胞增殖失控，继而形成肿瘤。因此，肿瘤被认为是体细胞遗传病。所有肿瘤都是基因突变的结果，这些基因参与控制细胞的正常生长过程、抑制细胞增殖或修复基因损伤等。

肿瘤的遗传基础十分复杂，某些基因的突变具有普遍性，与多种肿瘤的形成有关；另一些基因突变则只限于某种特定细胞类型。突变基因的不同组合可产生相同类型的肿瘤。体细胞内遗传物质的改变又与各种环境因素的直接或间接作用有关，在环境因素所引起的染色体或 DNA 损伤基础上，体细胞去分化并无限制地增殖形成肿瘤细胞，经过促进和进展等过程，最终形成肿瘤。

第一节　肿瘤的遗传易感性

不同人群、不同个体的遗传结构不同，呈现出肿瘤易感性的差异，表现在肿瘤发生率的种族差异、环境因素致癌的个体差异、肿瘤发生的家族聚集现象等方面。

一、肿瘤发病率的种族差异

遗传流行病学研究表明，不同种族间，同一种肿瘤的发病率存在显著差异。例如，日本人患松果体瘤比其他民族高 11 ~ 12 倍。再如，中国人鼻咽癌发病率居世界首位，中国人中又以广东人发病率最高，移居美国的华人鼻咽癌的发病率比美国白种人高 34 倍，这显然不能用环境因素来解释。种族差异主要是遗传差异，由此可见遗传因素在肿瘤发病中的作用。

二、肿瘤的家族聚集现象

大多数肿瘤在人群中是散发的，但不同家族的肿瘤发病情况不同，有些肿瘤或有些家族表现出肿瘤的家族聚集现象。

如果一个家族的几代中有多个成员患相同器官或不同器官的肿瘤，则称这样的家族为癌家族(cancer family)。癌家族有肿瘤发病年龄早、某类肿瘤(特别是腺癌)所占比例大、常染色体显性遗传等特征。典型的癌家族是 1913 年 Warthin 首报的 G 家族，其后多位学者相继对该家族进行了研究，研究成果记载了 1895 年至 1985 年该家族成员的肿瘤发病情况。在该家族 7 代 10 个支系 842 名后裔中，共有 95 名癌症患者，其中患结肠癌 48 人，子宫内膜癌 18 人；13 人患多发性肿瘤，19 人癌发生于 40 岁之前；72 名癌患者的双亲之一患癌，男患者与女患者之比为 47:48。发病情况符合癌家族的特征。1985 年，Lynch 等报道 18q11 ~ q12 的 *kidd* 血型基因与 G 家族肿瘤相关基因连锁。

如果一个家族多个成员患同一种类型的肿瘤，则该肿瘤称为家族性癌（familial carcinoma）。家族性癌的特点是：①一般为较常见的肿瘤，如乳腺癌、结肠癌、胃癌等；②家族中患者一级亲属发病率高于群体3～4倍；③同卵双生者发病一致率高；④遗传方式为多基因遗传。群体中，家庭性癌大多数仍然是散发的，少数有家族聚集性，例如，仅12%～15%的结肠癌患者有家族史。

三、环境因素致癌的个体差异

在人们的生活环境中存在着许多物理的、化学的和生物的致癌因子，在一定条件下，这些致癌因子可以诱发肿瘤。例如，各种电离辐射和紫外线照射可致皮肤癌和白血病；通过吸烟进入机体的3,4-苯并芘可以引起肺癌；黄曲霉素可以诱发肝癌；亚硝胺可以引起各种消化道肿瘤；某些病毒与鼻咽癌、白血病密切相关。这些致癌因子是通过改变遗传物质，导致基因突变进而引发肿瘤的。尽管人们都可能接触各种致癌因子，但并非人人都会发生癌症，这表明肿瘤的发生还与个体的易感性有关，而易感性在很大程度上是遗传因素决定的。

四、致癌因子代谢与肿瘤

环境中的致癌因子进入机体后的代谢过程，既可能使其失活进而排出体外，也可能使其活化而获得致癌性。直接致癌物或以原型与生物大分子结合，或自动降解活化为亲电子的终致癌物而发挥致癌性。体内的酶系统参与直接致癌物或终致癌物的失活过程。间接致癌物的原形分子没有损伤DNA能力，需经机体酶的活化作用才有致癌性，其失活也在生物转化酶系统催化下进行。酶的活性改变可以影响致癌因子在体内的代谢并促进肿瘤的发生。例如，芳烃羟化酶（AHH）是肝细胞中的一种氧化酶，能在体内活化许多致癌的多环芳烃，包括通过吸烟途径获得的各种芳烃，促进癌的发生；致癌性较弱的致癌因子，如苯蒽化合物，不仅能通过AHH的作用而活化，而且还有诱导AHH活性的作用；AHH的可诱导性在人群中呈多态性，可区分为高、中、低3个诱导组，据调查，如果以低诱导组患肺癌的易感性为1，则中诱导组为16，而高诱导组高达36。

五、免疫缺陷与肿瘤

人体内基因突变是经常发生的，其中少量突变细胞进入恶性细胞的始动过程，并形成恶性细胞，但人体的免疫系统通过T细胞的免疫监视作用，不断地将这些新产生的恶性细胞清除。因此，只要免疫监视功能正常，一般不会患肿瘤。然而，遗传性免疫缺陷能使突变细胞逃脱这种监视而发展成为肿瘤细胞。例如，呈X连锁隐性遗传的无丙种球蛋白血症（瑞士型）患者，不但免疫球蛋白缺乏，而且T细胞也缺乏，常因患病毒和真菌感染而早年夭亡，幸存的患者则易患白血病和淋巴系统肿瘤。再如，常染色体显性遗传的严重联合免疫缺陷症（SCID）患者易患白血病等造血系统肿瘤。除遗传性免疫缺陷外，艾滋病患者由于感染人类免疫缺陷病毒（HIV）而致继发性免疫缺陷，对肿瘤也有较高的易感性，尤其易患Kaposi肉瘤。

第二节　遗传性肿瘤与遗传性肿瘤综合征

遗传性肿瘤与遗传性肿瘤综合征都有明显的遗传倾向，受单基因控制，符合孟德尔遗传规律，大多数呈常染色体显性遗传。

一、遗传性肿瘤

遗传性肿瘤(hereditary tumor)大多起源于神经组织或早期胚胎组织，除呈现家族性、双侧性和多发性特征外，还常伴有先天性畸形及肿瘤抑制基因异常。

(一)视网膜母细胞瘤

视网膜母细胞瘤(retinoblastoma，RB)为儿童期发病的眼部肿瘤(图 11－1)，发病率约为 1/15 000，恶性程度高，分遗传型和非遗传型 2 类。遗传型约占 40%，遗传方式为 AD。

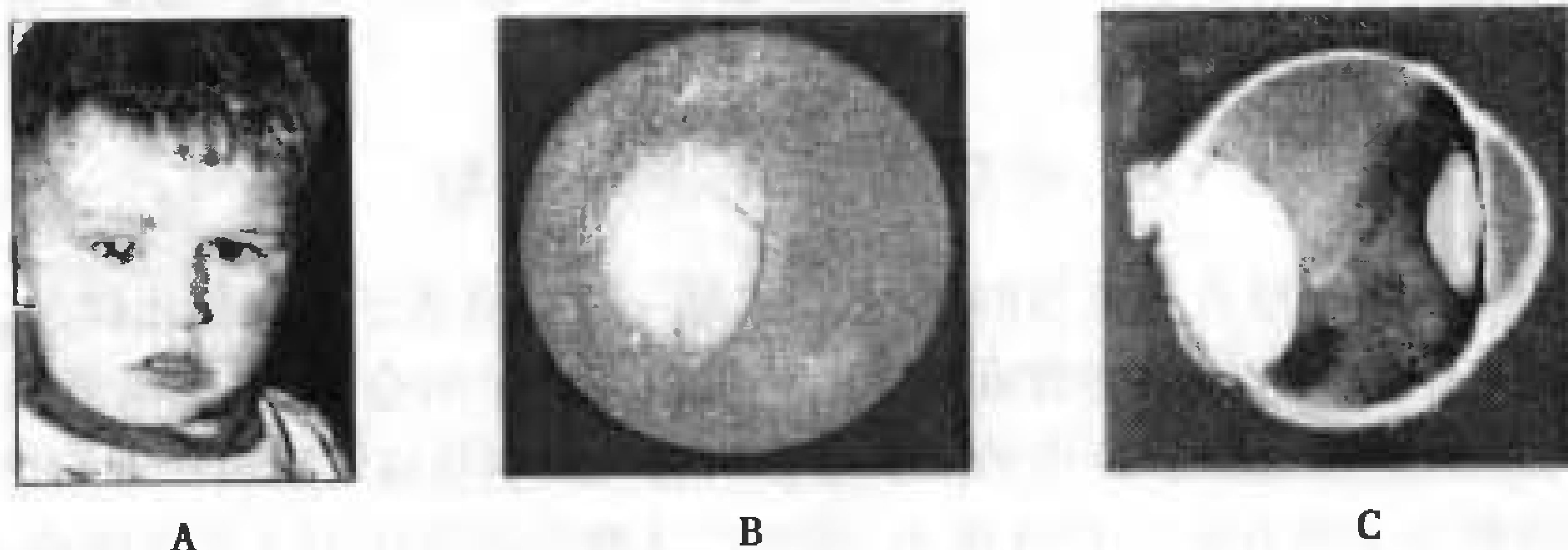


图 11－1　视网膜母细胞瘤

A：患儿外貌(猫眼)　B：早期表现　C.进展期表现

遗传型患儿的双亲之一携带突变的视网膜细胞瘤基因(*rb*)，或患儿父母生殖细胞基因突变，使患儿出生时全身细胞均含突变基因(*rb*)；出生后如视网膜细胞再发生 1 次基因突变(*rb*)，即可形成视网膜母细胞瘤(*rb*/*rb*)。遗传型患者常双眼发病，平均发病年龄为 15 个月。非遗传型患者的 2 次基因突变均发生在出生后，因此，常单眼发病，发病较迟，平均发病年龄为 30 个月。现已知视网膜母细胞瘤的发生是位于 13q14 的肿瘤抑制基因——视网膜母细胞瘤基因(*Rb*)突变所致。

(二)Wilms 瘤

Wilms 瘤也称肾母细胞瘤(nephroblastoma)，是一种婴幼儿期的肾肿瘤。大多在 4 岁前发病，活婴发病率约为 1/10 000，约占全部肾肿瘤的 6%。Wilms 瘤也分为遗传型(38%)和非遗传型(62%)2 类。遗传型多呈双侧性，发病年龄较早，遗传方式为 AD，有明显的家族聚集现象。患者可伴有无虹膜、假两性畸形、先天性心脏病以及智能低下等。

易患 Wilms 瘤的无虹膜症患者及 Wilms 瘤患者细胞中均发现 11p13 的缺失。近年来研究表明，Wilms 瘤(WT)基因是一种肿瘤抑制基因，定位于 11p13，患者的肿瘤组织中有 *WT* 基因的纯合缺失，其正常组织中则为杂合子，这说明其发病机理与视网膜母细胞瘤相同。

属遗传性肿瘤的还有黑色素瘤、神经母细胞瘤、嗜铬细胞瘤等。

二、遗传性肿瘤综合征

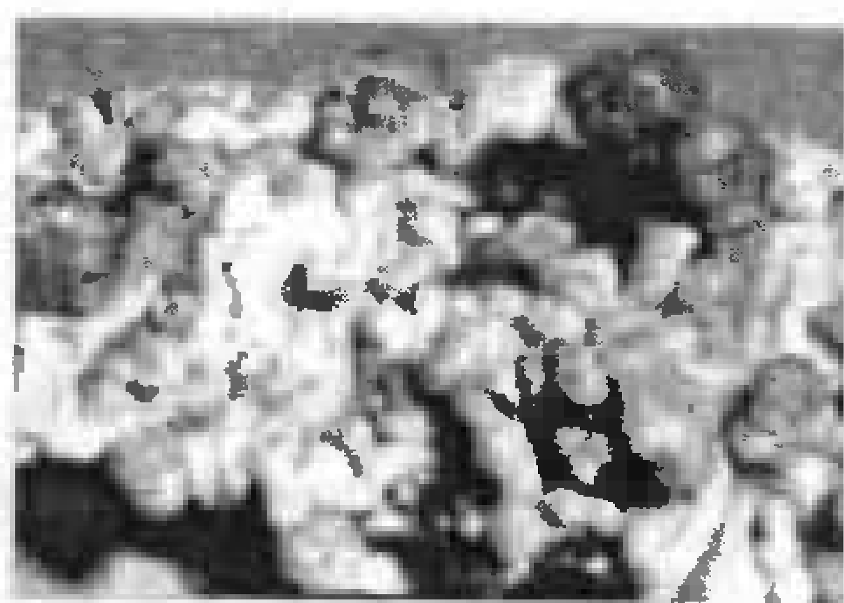
有些AD病晚期常并发肿瘤，也可以认为肿瘤是AD病的晚期表现，此类疾病称为遗传性肿瘤综合征(hereditary tumor syndrome)，也称为遗传性癌前病变。

(一)家族性结肠息肉

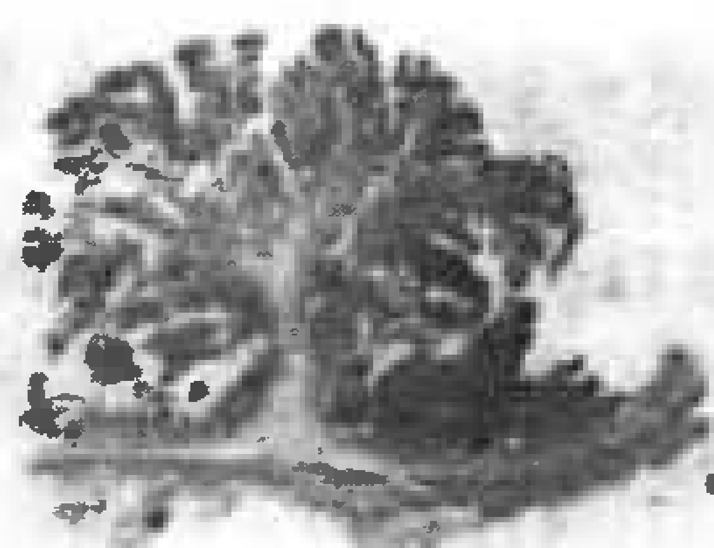
家族性结肠息肉(familial polyposis coli, FPC)又称家族性腺瘤样息肉，是一种AD病，本病具有家族性、遗传性、多发性和早发倾向。群体发病率约为1/10 000。患者在青少年时结肠和直肠已有多发性息肉，临床上多表现为肠梗阻或血性腹泻。随着息肉的长大、增多，大约在40岁时恶变为结肠癌(图11-2)。90%未经治疗的患者将死于结肠癌。该病的致病基因已定位于5p21~22。正常人此处存在肿瘤抑制基因 *APC*，FPC患者的致病基因实质上是有缺陷的 *APC* 基因。

(二)多发性神经纤维瘤

多发性神经纤维瘤(neurofibromatosis，NF1)患者躯干皮肤上分布有多发的神经纤维瘤(图11-3)，表现为多个浅棕色的"牛奶咖啡斑"，腋窝有雀斑。NF1可恶变为纤维肉瘤、鳞癌和神经纤维肉瘤。NF_1 基因也是肿瘤抑制基因，定位于17q11.2。



A



B

图11-2 家族性结肠息肉
A:结肠黏膜上布满息肉 B.结肠息肉恶变



图11-3 Ⅰ型神经纤维瘤
示背面多发性肿瘤

(三)基底细胞痣综合征

基底细胞痣综合征(basal cell nevus syndrome)也有遗传型和非遗传型之分。遗传型表现为多发性皮肤基底细胞痣，基底细胞痣可于青春期癌变；痣为多发性，可达近百个，病灶小，最小直径仅1 mm。基底细胞痣综合征基因定位于9q22~31。

第三节 染色体不稳定性综合征

某些AR病存在DNA修复系统缺陷，染色体具有不稳定性，易于断裂而重排，这类疾病称为染色体不稳定性综合征。染色体不稳定综合征临床特点是：至少对1种DNA损伤因子高度敏感，具有遗传不稳定性和肿瘤易感性，易患白血病或其他肿瘤。

一、Bloom 综合征

Bloom 综合征(Bloom syndrome,BS)患者身材矮小,发育迟缓,对日光敏感,面部有典型的蝴蝶斑,中度或重度免疫缺陷,多在 30 岁前发生各种肿瘤和白血病。体外培养淋巴细胞呈现各种类型的染色体畸变和染色单体畸变(包括对称的四射体),姐妹染色单体交换率(SCE)高于正常人约 10 倍。本病致病基因 *BLM* 基因定位于 15q26.1。正常 *BLM* 基因编码 RecQ DNA 解链酶家族的一员。由于 *BLM* 基因突变导致 RecQ DNA 解链酶活性缺乏,不能修复 DNA 复制过程出现的各种异常的 DNA 结构,导致染色体断裂、易位等。

二、Fanconi 贫血

Fanconi 贫血(fanconi anemia,FA)为儿童期的骨髓疾病,相当罕见,发病率约为 1/350 000,遗传方式为 AR。本病特点是进行性全血细胞减少,故又称先天性全血细胞减少症(congenital pancytopenia)。患者生长发育迟缓,有贫血、易疲乏、易出血和感染等临床表现。儿童期发生肿瘤的危险性增高,尤其易患白血病。患者染色体自发断裂率明显增高,单体断裂、裂隙等染色单体畸变多见;双着丝粒染色体、核内复制也很常见。FA 的基因比较复杂,目前已鉴定出 8 个 FA 互补组。已知 FAA 互补组的 *FAA* 基因定位于 16q24.3,FAC 互补组的 *FAC* 基因定位于 9q22、23。

三、共济失调性毛细血管扩张症

共济失调性毛细血管扩张症(ataxia telangiectasia,AT)遗传方式为 AR,临床特征为小脑性共济失调,眼毛细血管扩张(图 11－4),免疫功能缺陷导致肺部反复感染及肿瘤易感性高,本病多见于儿童期,1 岁左右即可发病,6 岁以后眼、面部和颈部出现瘤样小血管扩张。AT 患者对电离辐射敏感,染色体畸变是非随机的,常见 14/14 易位或其他涉及 14 号染色体的改变,14q12 为断裂热点。AT 患者易患多种肿瘤,特别是白血病、淋巴瘤,肿瘤发病率高出正常人群 100 余倍。经过近 40 年的研究,AT 的致病基因 *ATM* 终于被克隆,该基因定位于 11q22。

四、着色性干皮病

着色性干皮病(xeroderma pigmentosum,XP)是一种罕见的致死性 AR 病,主要临床特点为早发的起源于表皮鳞状细胞或基底层细胞的皮肤癌。其他临床表现还包括性发育障碍、生长发育迟缓、智能低下、神经性耳聋等。患者皮肤对紫外线特别敏感,易出现皮疹和色素沉着,皮疹和色素沉着处常常是皮肤癌的好发部位(图 11－5)。患者染色体自发断裂的频率不高,但紫外线照射后染色体断裂频率明显升高。

与 XP 相关的基因很多,绝大多数已知的 *XP* 基因突变使体内核苷酸切除修复系统失去功能,不能有效切除紫外线诱发的嘧啶二聚体,继而导致突变率增高,形成肿瘤。



图 11－4　共济失调性毛细血管扩张症(示眼结膜毛细血管扩张)



图 11－5　着色性干皮病
A:面部红斑　B.面部皮肤癌变

第四节　染色体异常与肿瘤

细胞内染色体数量和结构是稳定的,染色体损伤是肿瘤发生的基础,而癌变后染色体稳定性降低又进一步导致染色体结构变化。因此,染色体异常是肿瘤细胞的重要特征之一。

染色体病患者除具有相应的多种临床表现外,其肿瘤发病率比一般人群高数倍乃至数十倍。有资料显示:21 三体综合征患者急性白血病的发病率比正常人高 15～18 倍,Klineflter 综合征患者易患男性乳腺癌,Turner 综合征患者条索状卵巢易恶变为卵巢癌,两性畸形患者发育不全的性腺(残留的睾丸组织)易发生精原细胞瘤。

一、肿瘤细胞的染色体畸变

(一) 染色体数目异常

肿瘤细胞多数为非整倍体,如亚三倍体、超三倍体、亚四倍体等。肿瘤细胞染色体数目的增多或减少并不是随机的。例如,许多肿瘤常见 8、9、12 和 21 号染色体的增多或 7、22、Y 染色体的减少;实体瘤的染色体数目变化一般在二倍体上下,或在三倍体和四倍体之间,一般属继发性改变;癌性胸、腹腔积液的染色体数目变化更大,可见上百条染色体。染色体数目变化的程度并不与肿瘤的恶性程度成正比。比如,胃癌细胞一般只有 1～2 条染色体的改变,但其恶性程度却很高。另外,同一肿瘤中,各肿瘤细胞的染色体数也不完全相同,甚至差别很大,但大多数肿瘤都可以见到 1～2 个干系,干系细胞的百分比不固定。

(二)染色体结构异常

肿瘤细胞染色体结构异常包括易位、缺失、重复、环状染色体和双着丝粒染色体等。在某种肿瘤或某类肿瘤中恒定出现,具有特定形态的结构异常染色体,称为特异性标记染色体(marker chromosome)。特异性标记染色体的存在支持肿瘤单克隆起源理论。目前发现的具有临床应用价值的特异性标记染色体是慢性粒细胞性白血病的 Ph 染色体和 Burkitt 淋巴瘤的 $14q^{+}$染色体。

1. Ph 染色体　1960 年，Nowell 和 Hungerford 在美国费城首先发现慢性粒细胞性白血病(CML)患者的骨髓及外周血淋巴细胞中有 1 个比 G 组染色体还小的染色体，命名为 Ph 染色体。最初认为，Ph 染色体是 22 号染色体的长臂缺失所致，后经显带证明是 9 号和 22 号染色体长臂相互易位所致(图 11－6)。易位后形成 2 条衍生染色体，一条是 $9q^+$ 染色体(9pter→9q34::22q11→22qter)，另一条是 $22q^-$ 染色体(22pter→22q11::9q34→9qter)，即 Ph 染色体。Ph 染色体的重要临床意义在于：大约 95% 的 CML 病例都是 Ph 染色体阳性，因此可以作为诊断的依据；与 CML 有相似临床表现的其他血液病，如骨髓纤维化等，Ph 染色体为阴性，故 Ph 染色体可用作鉴别诊断；Ph 染色体往往先于临床症状出现，故又可用于早期诊断。此外，Ph 染色体阴性的 CML 患者对治疗反映差，预后不佳。

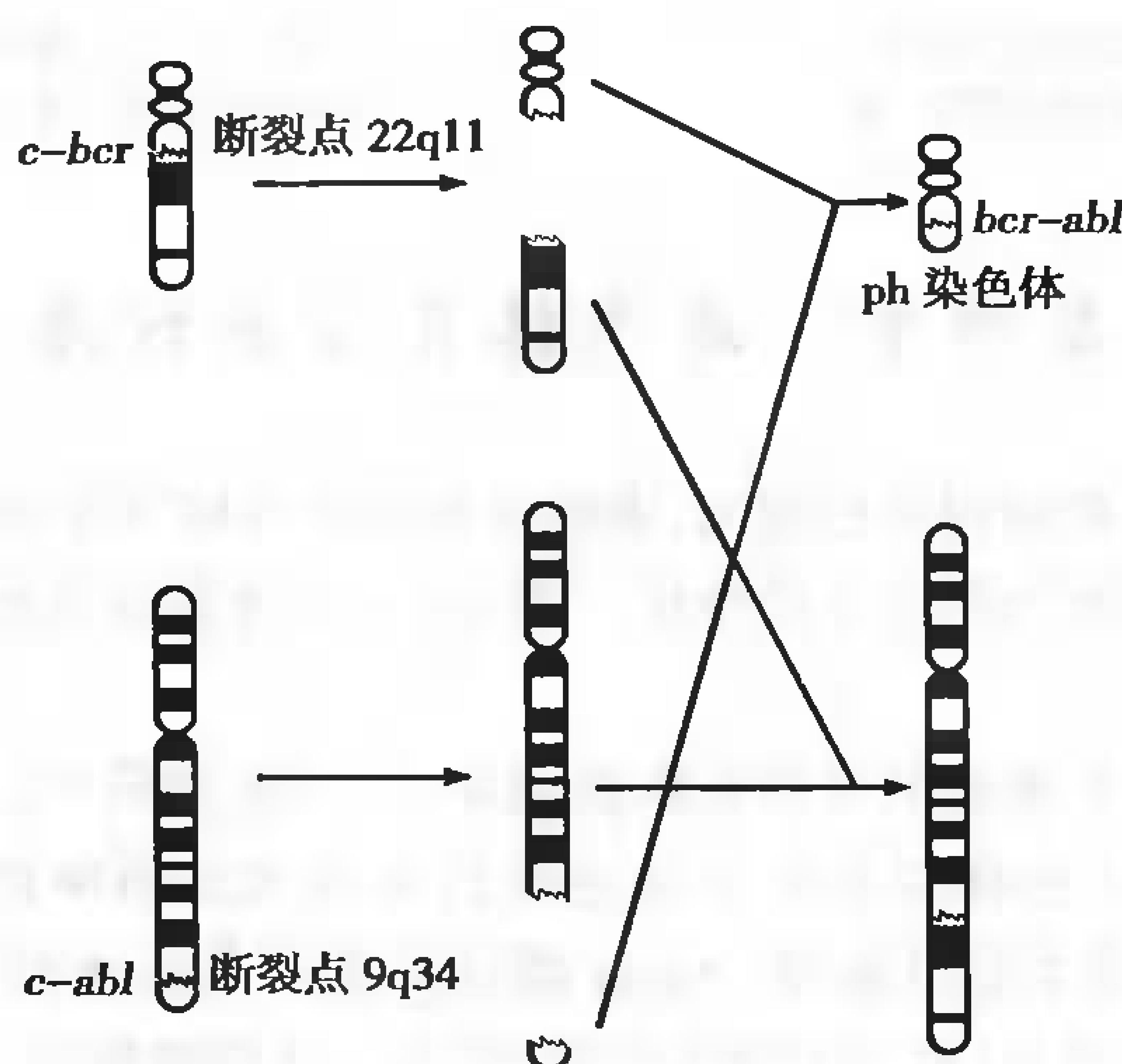


图 11－6　CML 患者的 9/22 易位(示 Ph 染色体)

2. $14q^+$ 染色体　$14q^+$ 染色体见于 75% 的 Burkitt 淋巴瘤患者，是 8号染色体长臂和 14 号染色体长臂相互易位形成的(图 11－7)。易位后形成 2 条衍生染色体：一条是 $8q^-$ 染色体(8pter→8q24::14q32→14qter)；另一条是 $14q^+$ 染色体(14pter→14q32::8q24→8qter)。

除 Ph 染色体和 $14q^+$ 染色体外，与特定肿瘤相关的非随机存在的特异性标记染色体还有不少，但检出频率均不高，这些标记染色体包括：脑膜瘤中的 $22q^-$ 或 －22，视网膜母细胞瘤中的 del(13)(p14)，小细胞肺癌的 del(3)(p14p23)，黑色素瘤的 ＋7 或 ＋22，鼻咽癌的 t(1;3)(q41;p11)等。

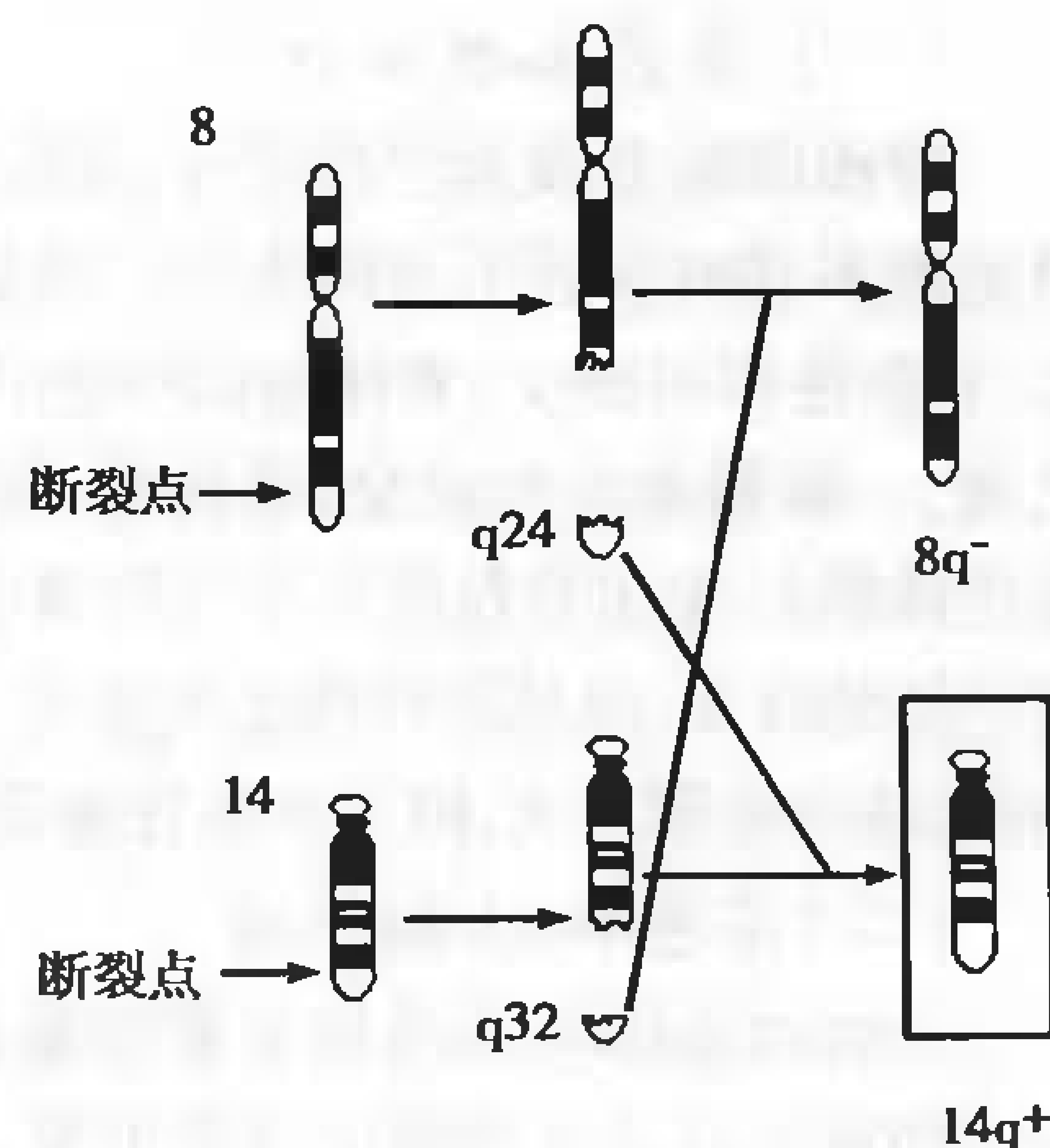


图 11－7　Burkitt 淋巴瘤患者的 8/14 易位(示 $14q^+$ 染色体)

二、染色体脆性部位与肿瘤

在接触某种特殊化学物质或在体外细胞培养条件下，人类染色体上某些特异部位表现出非随机的断裂或裂隙，染色体上这些特异部位称为脆性部位(fragile sites, fra)。脆性部位的存在是染色体不稳定性的体现。截止HGM10的资料，已有113个脆性部位在染色体上被确认，除21号染色体外，每条染色体上均有脆性部位。按遗传特性和表达频率，脆性部位分为罕见型和常见型两类。

某些肿瘤的染色体畸变断裂点与罕见型脆性部位一致。例如，Burkitt淋巴瘤患者染色体t(12;14)(q13;q32)易位的断裂点之一与罕见型脆性部位12q13一致，急性粒细胞性白血病患者染色体inv(16)(p13q32)倒位的断裂点之一与罕见型脆性部位16q32一致。学者们认为，这些脆性部位的表达为染色体断裂和重排提供了条件，而断裂和重排则是癌变的基础。

常见型脆性部位通常表达频率很低，但在某些肿瘤患者则有高表达。例如，Wilms瘤患者有11号染色体缺失del(11)(p13p14)，小细胞肺癌患者有3号染色体短臂缺失del(3)(p14p23)，这2种缺失分别涉及常见型脆性部位11p13、3p14，这说明常见型脆性部位表达频率增高是染色体稳定性降低的标志，在此基础上发生的各种染色体重排将导致肿瘤发生。

第五节　肿瘤相关基因

基因改变是肿瘤起源和发展的分子基础，这些基因统称为肿瘤相关基因。主要包括：癌基因、肿瘤抑制基因、肿瘤转移基因和肿瘤转移抑制基因等。

一、癌基因

癌基因(oncogene)是指能使细胞癌变的基因，包括存在于病毒基因组的病毒癌基因和存在于细胞核基因组的细胞癌基因。

(一)病毒癌基因和细胞癌基因

可引起动物肿瘤的逆转录病毒不能自我扩增，为缺陷型逆转录病毒，其基因组有与病毒基因序列不同的高度保守的插入序列，杂交分析表明，插入序列来源于宿主细胞基因组；进一步研究表明，插入序列来源于宿主细胞的RNA分子。当病毒感染宿主细胞后，插入序列编码的蛋白在宿主细胞中成为肿瘤的强诱导剂，并且插入序列可以在逆转录后整合到宿主细胞的基因中，经宿主细胞翻译产生的插入序列编码的蛋白，具有解除细胞增殖限制的作用，进而促进肿瘤的发生。这种具有引发肿瘤作用的逆转录病毒基因组中的插入序列称为病毒癌基因(viral oncogene, v－onc)。

细胞癌基因(cellular oncogene, c－onc)也称原癌基因(proto－oncogene)，是胚胎生长发育必需的一类基因，出生后大多已关闭，不再表达。一旦这些基因在表达时间、表达部位、表达数量及表达产物结构等方面发生了异常，就可导致细胞无限制增殖并形成肿瘤。细胞癌基因的转录产物经剪切加工后形成的RNA分子，如被逆转录病毒俘获并插入到逆转

录病毒基因组中，则成为病毒癌基因。这说明细胞癌基因与病毒癌基因具有同源性，二者的主要的差别是病毒癌基因中不含内含子，外显子部分也可能存在基因突变(图 11-8)。

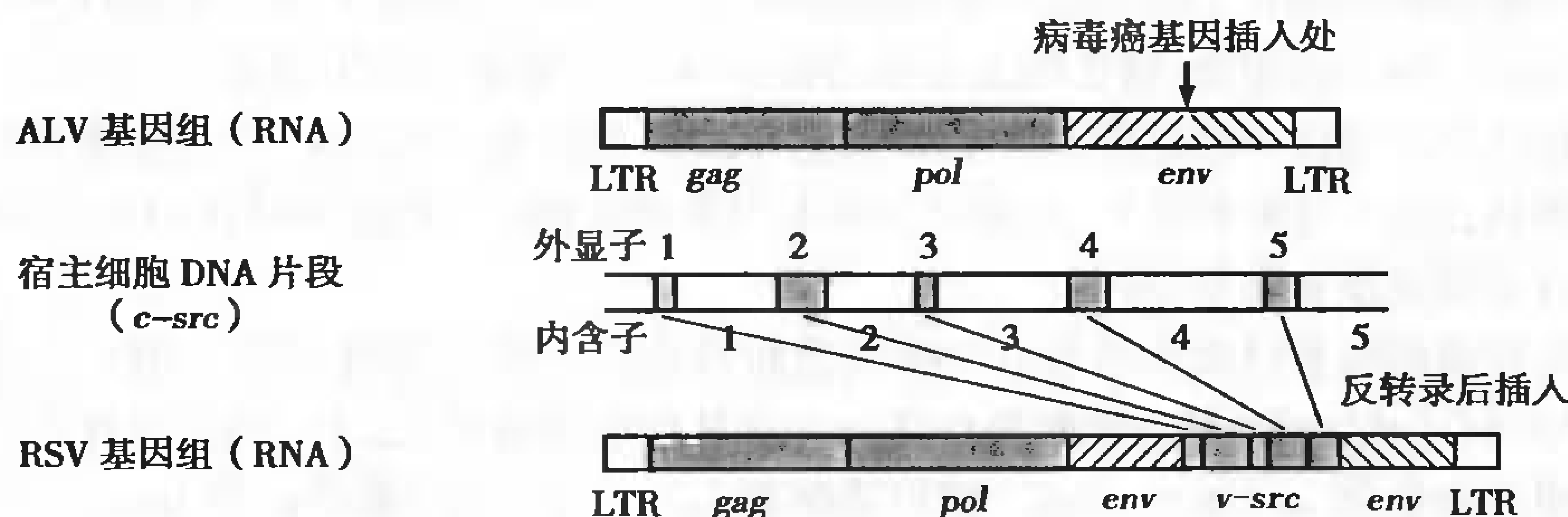


图 11-8　病毒癌基因和细胞癌基因

LTR:长末端重复序列　ALV:禽类白血病病毒,RSV 的前身　RSV:Rous 肉瘤病毒

(二)癌基因及其产物的命名

病毒癌基因以首次发现能诱发动物肿瘤的该基因所在的逆转录病毒的缩写来表示，并冠以“v”字，表示来源于病毒基因组。例如，首次发现的能诱发猿猴形成肉瘤的猿猴肉瘤病毒(simian sarcoma virus)的病毒癌基因命名为 *v - sis*。与病毒癌基因同源的细胞癌基因的命名，沿用病毒癌基因的名称，有时为了同病毒基因相区别，细胞癌基因也可用大写字母表示，并冠以“C”字，如 *C - MYC*。有的病毒癌基因家族有多个成员，命名时分别加上相应的前缀，如在 Harvey 鼠肉瘤病毒和 Kirsten 鼠肉瘤病毒中发现的 *ras* 基因分别表示为 *v - Ha - ras*，*v - Ki - ras*。癌基因产物的命名主要依据相对分子质量，有时也兼顾蛋白质的性质，如 3 种 *ras* 基因的产物均为 21 kDa 的蛋白质，表示为 $p21^{ras}$；*v - src* 基因的产物为 60 kDa 的磷蛋白，表示为 $pp60^{v-src}$，其中第 1 个“p”代表蛋白质，第 2 个“p”代表磷蛋白。

(三)细胞癌基因的分类

目前，已发现人类细胞癌基因 100 余种，同时确定了它们在染色体上的位置。这些基因与细胞生长、增殖和细胞周期调控等功能有关。它们有的编码生长因子、生长因子受体或蛋白激酶，在生长信号的传递和细胞分裂中发挥作用；有的编码 DNA 结合蛋白，参与基因的表达或复制的调控。根据基因的产物不同，细胞癌基因分为 4 类：①蛋白激酶类，此类细胞癌基因的产物具有受体功能，各种生长因子与相应受体结合，能引起受体酪氨酸激酶活化，引起细胞增殖；②生长因子类，该类细胞癌基因产物为生长因子，能与相应生长因子受体结合并引起细胞增殖；③细胞内信号转导蛋白类，该类细胞癌基因的产物为在细胞内起信号转导作用的 G 蛋白类。生长因子与相应的受体结合后，要通过细胞内的一系列分子间的相互作用，才能最终把信号传至核内，引起细胞增殖；④核内转录因子类，此类细胞癌基因的产物为 DNA 结合蛋白，参与基因的表达和复制的调控。

(四)细胞癌基因的激活机制

不同的细胞癌基因通过不同机制激活。一对细胞癌基因中只要有 1 个被激活，就可以以显性方式发挥作用，使原来处于关闭状态的细胞癌基因表达基因产物，或者表达过量的产物或异常的产物，从而使正常细胞恶变(图 11-9)。细胞癌基因的激活机制有 4 种。

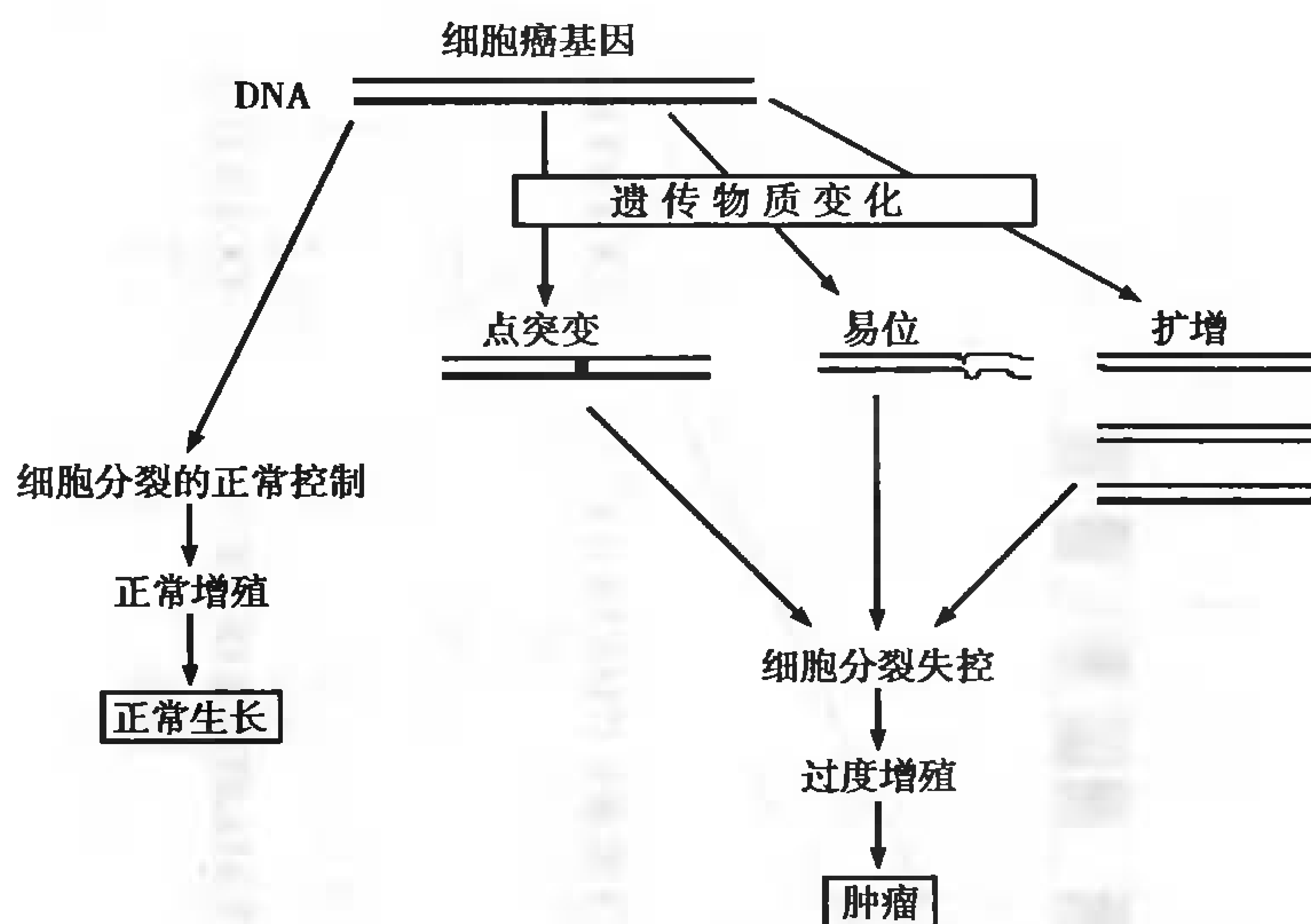


图 11-9 细胞癌基因的激活机制

1.病毒诱导与启动子插入 *c*-*onc* 附近一旦插入一个强大的启动子,就可被激活。如逆转录病毒感染细胞后,其基因组两端的长末段重复序列(LTR)插入到 *c*-*onc* 的适当位置,其启动子可使 *c*-*onc* 激活,这是动物细胞 *c*-*onc* 的主要激活方式。带有 LTR 的禽类白细胞增生病毒(ALV)整合到 *c*-*myc* 附近,可使 *c*-*myc* 表达水平高于正常细胞 30~50 倍。

2.点突变 点突变使 *c*-*onc* 编码的基因产物异常,也可使基因脱离正常调控而过度表达。例如,膀胱癌细胞株中 *c*-*ras* 的 12 位密码子 GGC 变为 GTC,使甘氨酸变为缬氨酸,结果导致细胞具有转化细胞的特征。

3.染色体断裂与重排 由于染色体断裂与重排,导致 *c*-*onc* 在染色体上的位置改变,使 *c*-*onc* 激活或形成异常蛋白而使细胞转化。

Burkitt 淋巴瘤是染色体易位激活 *c*-*onc* 的典型例子。在 Burkittt 淋巴瘤患者中,75%为 t(8;14)(q24;q32);16%为 t(8;22)(q24;q11);9%为 t(2;8)(p12;q24)。这 3 种易位均涉及 8q24 *c*-*myc*;参与易位的另外 3 条染色体的断裂点在免疫球蛋白重链基因 *IGH*(14q34)或轻链基因 *IGL*(22q11)、*IGK*(2p12)处(图 11-10)。易位使位于 8q24 的 *c*-*myc* 分别受 *IGH*、*IGL*、*IGK* 转录调节区启动子、增强子的影响而活化,转录活性明显增高,过量表达的 *c*-*myc* 产物导致细胞恶变。

绝大部分 CML 患者存在 t(9;22)(q34;q11)易位,9q34 处存在 *c*-*abl*,22q11 处存在 *c*-*bcr*,带有细胞癌基因 *c*-*abl* 的 9q34→qter 易位于 22q11 的 c-bcr 处形成 bcr→abl 融合基因(图 11-6)。正常的 *c*-*abl* 的产物编码相对分子质量为 145 kDa 的酪氨酸激酶(p145),只有低活性,且受"生长因子-受体"系统的调控;*abl*-*bcr* 融合基因产物 p210 表现为增强的酪氨酸激酶活性,并且脱离"生长因子-受体"系统的调控,引起细胞癌变。

4.基因扩增 此种激活方式是指 *c*-*onc* 的结构本身没有发生改变,但其拷贝数由于

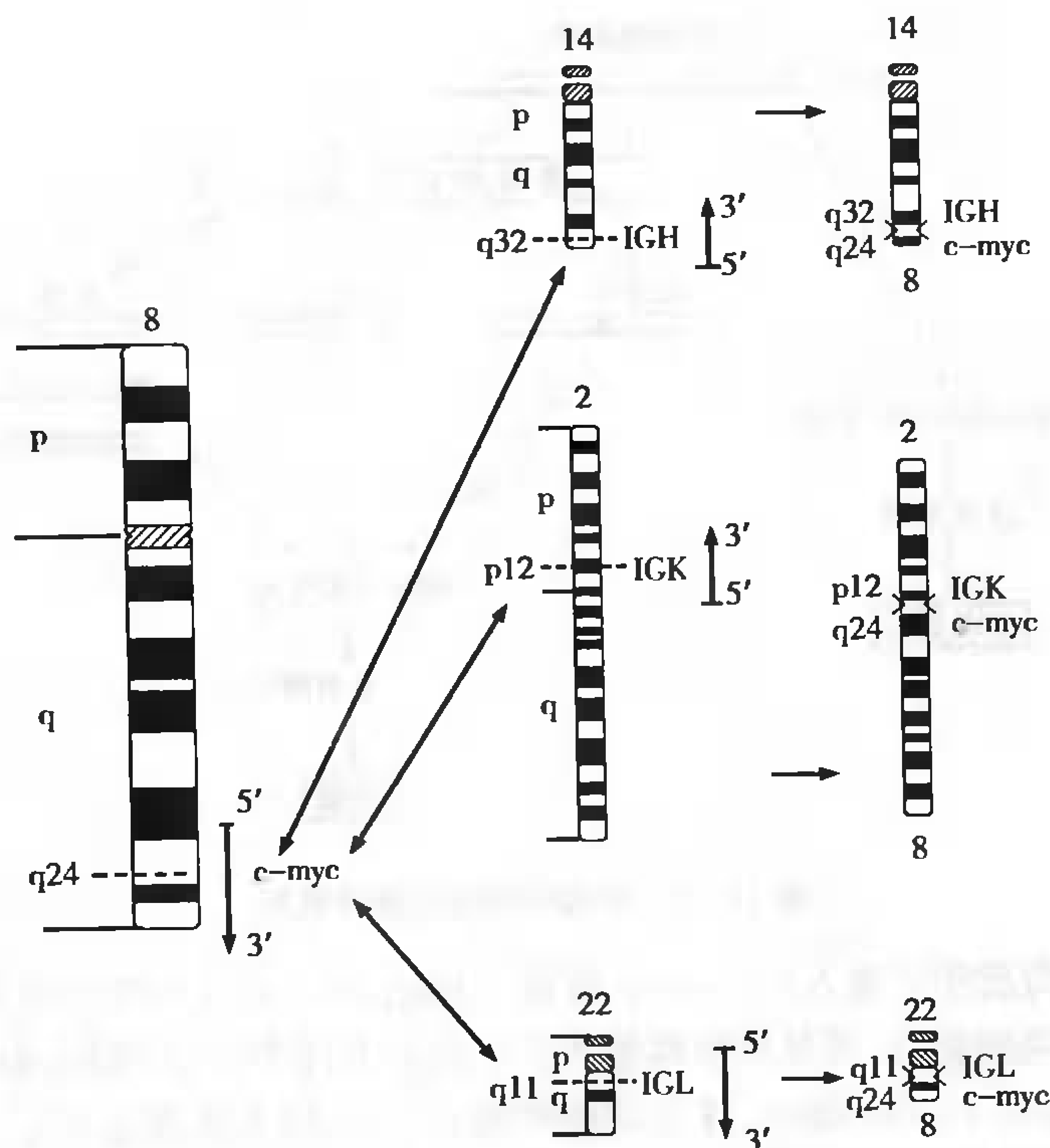


图 11-10 Burkitt 淋巴瘤患者染色体易位导致 c-myc 活化

DNA 不断复制而增多，导致 *c-onc* 产物过量表达。在细胞水平，这种扩增表现为双微体(double minutes, DMs)和染色体上的均染区(homogeneously staining region, HSR)。DMs 是扩增的 DNA 脱离染色体而形成的无着丝粒的微小染色体，成对存在；HSR 是扩增的 DNA 在染色体的某些节段上呈现的、相对解旋的浅染区。研究表明，*c-onc* 的扩增往往出现在肿瘤的进展阶段，因此与预后密切相关，扩增的拷贝数越多，预后越差。神经母细胞瘤的 *c-N-myc* 扩增可以是正常细胞的 140～250 倍。

二、肿瘤抑制基因

肿瘤抑制基因(tumor suppressor gene)也称抗癌基因(anti-oncogene)，是 20 世纪 60 年代在肿瘤细胞与正常细胞杂交研究的基础上提出的。正常细胞与肿瘤细胞融合形成的杂种细胞没有肿瘤细胞表型，正常细胞的染色体可以逆转肿瘤细胞的表型。因此，人们认为正常细胞中可能存在抑制肿瘤发生、促进细胞分化的基因，即肿瘤抑制基因。1988 年，首次在人类视网膜母细胞瘤中发现的 *RB* 基因，证实了当初人们对肿瘤抑制基因的推测，至今已确认了十几种肿瘤抑制基因。

正常细胞中某肿瘤抑制基因的 2 个等位基因，其中 1 个因突变或缺失而丧失功能，细胞仍具有抑制肿瘤形成的作用；如果 2 个都丧失功能，即处于纯合失活状态时，细胞就会因正常抑制作用的解除而恶性转化。因此，肿瘤抑制基因也称为隐性癌基因(recessive

oncogene)，即致癌作用是隐性的。少数肿瘤抑制基因存在例外的情况，例如，肿瘤抑制基因 *APC* 表现为显性遗传模式，在家族性结肠息肉的癌变过程中，一个 *APC* 等位基因的丢失发生在结肠腺瘤阶段，另一个丢失发生在结肠腺瘤向腺癌转变的阶段。以下简介最重要的 2 个肿瘤抑制基因 *Rb* 基因和 *p*53 基因。

Rb 基因定位于 13q14.1，全长约 200 kb，含 27 个外显子，编码 928 个氨基酸的核磷酸蛋白质(prb)，相对分子质量为 110 kDa。prb 通过抑制转录因子 E2F 而抑制细胞周期，因此具有抑制肿瘤细胞增殖的作用。

遗传型视网膜母细胞瘤患者出生时，由于亲代的 1 个 *Rb* 基因丧失功能或亲代生殖细胞发生 *Rb* 基因突变，导致患者体内所有细胞均为 *Rb* 基因的杂合状态，但由于 *Rb* 基因的致癌作用是隐性的，故不发病；出生后，如患者视网膜细胞中另一个 *Rb* 基因突变，则导致 *Rb* 基因纯合失活形成肿瘤。*Rb* 基因的缺失或功能丧失除见于视网膜母细胞瘤外，还见于骨肉瘤、小细胞肺癌、乳腺癌等。

*p*53 基因因其基因产物相对分子质量为 53 kDa 而得名，该基因定位于 17p13.1，全长 20 kb，含 11 个外显子，基因产物 *p*53 蛋白是核内的一种磷酸化蛋白质，含 393 个氨基酸。DNA 损伤、应激等可引起细胞内野生型 *p*53 蛋白水平升高，大量的野生型 *p*53 蛋白作为转录因子与特异的 DNA 序列结合，激活一系列下游靶基因的转录，进而诱导细胞周期 G_1 期阻断、细胞凋亡和细胞分化，保护基因组的完整性及抑制肿瘤细胞的生长等。*p*53 基因产生突变，细胞将失去这些功能，导致肿瘤发生。根据目前的研究，*p*53 基因的抑癌作用具有“广谱”性，在 50%左右的人类肿瘤中存在 *p*53 基因异常。

三、肿瘤转移基因与肿瘤转移抑制基因

肿瘤的侵袭和转移是肿瘤与宿主细胞、细胞外基质之间一系列复杂的、多步骤、多因素相互作用的动态过程。近年来，在肿瘤转移机制的研究中发现了一些与肿瘤转移呈正、负相关的肿瘤转移基因和肿瘤转移抑制基因。

(一)肿瘤转移基因

肿瘤转移基因(tumor metastatic gene)是指基因的表达和改变能够促进和导致肿瘤细胞发生转移的基因。人们曾经确信，肿瘤的致癌性和转移性各有一套基因决定，但后来发现某些癌基因转染适宜的受体细胞可引起转移表型的发生，表明已知的肿瘤形成相关基因，包括癌基因和突变型肿瘤抑制基因，具有肿瘤转移效应。一些编码细胞表面受体的基因也与肿瘤细胞的转移有关。整合素(integrins)是一类细胞表面粘合受体，能识别细胞基质中的黏附蛋白，起着固定细胞，抑制其转移的作用。这些受体基因的突变和功能丧失对肿瘤细胞的转移有促进作用。

(二)肿瘤转移抑制基因

肿瘤转移抑制基因(tumor metastasis suppressor gene)是指能抑制肿瘤转移形成的基因。例如，金属蛋白酶组织抑制因子基因编码的糖蛋白能与胶原酶结合，抑制其活性，保护基底膜免受胶原酶的降解，而基底膜是阻止肿瘤转移的主要屏障。因此，金属蛋白酶组织抑制因子基因属肿瘤转移抑制基因。

目前，nm23 被认为是肿瘤转移抑制基因的最佳代表。*nm*23 基因定位于 17q12 – q21，

编码的分子量为 17 kDa 蛋白质可与 G 蛋白结合，有影响微管生成、进而抑制肿瘤细胞转移的作用。动物肿瘤模型显示 nm23 RNA 在转移性肿瘤细胞中下调，在多种人类肿瘤，如乳腺癌、肾癌、黑色素瘤、肺癌、结肠癌等的肿瘤组织中，*nm23* 的表达与患者的预后及肿瘤的淋巴结转移有关。

第六节 肿瘤发生的遗传学说

一、单克隆起源学说

肿瘤是体细胞突变的结果，同一肿瘤的不同肿瘤细胞，染色体异常具有相似性，这表明他们都来源于一个共同的突变细胞，经过分裂而形成克隆，即肿瘤起源于单个细胞（单克隆）。但是肿瘤细胞群体又受到环境的影响，处于不断的变异之中。因此，这些肿瘤细胞的核型常常不完全相同，而且在同一肿瘤的发展过程中，核型也在不断演变。一些染色体畸变是致死的，而另一些畸变却能使细胞获得生长优势，因此肿瘤细胞群体经常处于变异和选择之中，此过程称为克隆演化（clone evolution）。肿瘤细胞群通过变异和选择，逐渐形成占主导地位的细胞群体，即干系（stem line）。干系即该肿瘤具有最常见核型的细胞群体，肿瘤的生长主要是干系的增殖。干系的染色体数称为众数（modal number）。有的肿瘤没有明显的干系，有的则可以有 2 个或 2 个以上的干系。干系以外的非主导细胞系称为旁系（side line）。由于环境条件的改变，干系可以转变为旁系，旁系可以发展成为干系。

二、二次突变学说

20 世纪 70 年代，Knudson 在分析同一类肿瘤中遗传型和非遗传型之间的关系时，提出了二次突变学说（two mutation theory）。他以遗传型和非遗传型视网膜母细胞瘤为研究对象，并假设 2 种类型视网膜母细胞瘤都是由 2 个独立与连续的基因突变产生的，即二次突变事件引起的。遗传型病例中，第 1 次突变发生于生殖细胞或亲代携带突变基因，并且传递给胚胎发育的每一个体细胞，而第 2 次突变随机发生在体细胞中。在这种情况下，双侧视网膜的细胞都有可能发生第 2 次突变并形成肿瘤。相比之下，非遗传型视网膜母细胞瘤是同一个体细胞发生 2 次独立突变形成的，因而双侧视网膜都发生二次突变的可能性较小。

根据二次突变学说，肿瘤发生是一种隐性事件，即野生型基因产物可以抑制肿瘤形成，而肿瘤中这一对等位基因均失活。当时 Knudson 称这种基因为肿瘤抑制基因，后来 *RB* 基因被发现并被克隆，以及其他肿瘤抑制基因被相继发现，证实了 Knudson 二次突变学说的合理性。

三、多步骤损伤学说

目前的研究证明，肿瘤的发生是多步骤、涉及多种相关基因、包括癌基因和肿瘤抑制基因变异的过程。一种肿瘤会有多种基因的变化，而同一种基因的改变也会在不同种类肿瘤的发生中起作用，大多数肿瘤的发生与癌基因的活化和（或）肿瘤抑制基因的失活有关。

目前认为，细胞癌变往往需要多个肿瘤相关基因的协同作用，要经过多阶段的演变，其中不同阶段涉及不同的肿瘤相关基因的激活或失活。这些基因的激活或失活在时间上有先后顺序，在空间位置上也有一定的配合，所以肿瘤细胞表型的最终形成是这些被激活或失活的相关基因共同作用的结果。在肿瘤的起始阶段，细胞癌基因激活的方式主要表现为逆转录病毒的插入和细胞癌基因点突变，而染色体重排、基因重组和基因扩增等激活方式的出现则意味着肿瘤进入演进阶段。正是由于各种细胞癌基因发生了质变和量变，导致表达异常，造成细胞分裂与分化失控，通过多步骤演变为肿瘤细胞。在多步骤演进过程中，除细胞癌基因激活外，还包括病毒癌基因的整合、肿瘤抑制基因的突变或缺失等，这些都是多步骤致癌过程的重要环节。

小　结

肿瘤遗传学是研究肿瘤发生与遗传因素之间关系的医学遗传学分支学科。肿瘤细胞是积累了不同基因突变的体细胞，突变导致细胞增殖失控，继而形成肿瘤。

不同人群或不同个体的遗传结构不同，呈现出肿瘤易感性的差异，表现在肿瘤发生率的种族差异、环境因素致癌的个体差异、肿瘤发生的家族聚集性现象等方面。

遗传性肿瘤与遗传性肿瘤综合征都有明显的遗传倾向，受单基因控制，符合孟德尔遗传规律，大多数呈常染色体显性遗传。遗传性肿瘤包括视网膜母细胞瘤、Wilms 瘤等；遗传性肿瘤综合征也称为遗传性癌前病变，包括家族性结肠息肉症、Ⅰ型神经纤维瘤和基底细胞痣综合征等。染色体不稳定综合征临床特点是：至少对 1 种 DNA 损伤因子高度敏感，具有遗传不稳定性和肿瘤易感性，易患白血病或其他肿瘤；包括 Bloom 综合征、Fanconi 贫血、共济失调性毛细血管扩张症和着色性干皮病。

染色体损伤是肿瘤发生的基础，染色体异常是肿瘤细胞的重要特征之一。在某种肿瘤或某类肿瘤中恒定出现，具有特定形态的结构异常染色体称为特异性标记染色体。Ph 染色体和 $14q^{+}$ 染色体是临床运用价值较大的 2 种特异性标记染色体。

基因改变是肿瘤起源和发展的分子基础，这些基因统称为肿瘤相关基因，主要包括：癌基因、肿瘤抑制基因、肿瘤转移基因和肿瘤转移抑制基因等。

癌基因（oncogene）是指能使细胞癌变的基因，包括存在于病毒基因组的病毒癌基因和存在细胞核基因组的细胞癌基因。细胞癌基因分为 4 类，即蛋白激酶类、生长因子类、细胞内信号转导蛋白类和核内转录因子类。细胞癌基因的激活机制有 4 种，即病毒诱导与启动子插入、点突变、染色体断裂与重排、基因扩增。

肿瘤抑制基因也称抗癌基因，致癌作用是隐性的，Rb 基因和 p53 基因是 2 种重要的肿瘤抑制基因。肿瘤转移基因是指基因的表达和改变能够促进和导致肿瘤细胞发生转移的基因，肿瘤转移抑制基因是指能抑制肿瘤转移形成的基因。

肿瘤发生的遗传学说包括单克隆起源学说、两次突变学说和多步骤损伤学说等。

（贡昌春　蔡绍京）

第十二章　药物遗传学

药物遗传学又称药理遗传学，是药理学与遗传学相结合发展起来的边缘学科，属于生化遗传学的范畴。它主要研究机体的遗传因素对药物代谢和药物反应的影响，特别是遗传因素引起的异常药物反应。

早在19世纪后半叶，人们就已经认识到了药物代谢、药物反应与遗传的关系。而正式将药物遗传学作为医学遗传学的一门分支学科提出来的是Motulsky和Vogel。1957年，Motulsky提出遗传决定的酶异常可引起药物的异常反应，Vogel则于1959年首次提出并使用“药物遗传学”这一术语。20世纪50年代以来，药物遗传学得到了迅速发展。研究方法由传统的双生子研究、系谱研究等发展到了现代的基因克隆、重组DNA等研究技术；研究内容由主要从单基因角度对药物与遗传关系的研究发展到从基因组整体水平研究药物与遗传的关系。做为医学生，掌握药物遗传学的基本规律对研究医学、药理学和人类遗传学的基础理论有重要意义，也对按照用药个体化原则正确用药、提高药效、减少或减轻不良反应也具有重要意义。

第一节　药物反应的遗传基础

药物作用受多种因素的影响，使药物代谢与药物反应出现个体差异。这些因素包括药物的化学结构、给药途径、个体的年龄、性别、营养状态、病理状态、遗传因素等。研究发现，同卵双生子(MZ)之间，异巴比妥、安替比林、双香豆素、水杨酸钠等药物的代谢动力学几乎没有差异，而异卵双生子(DZ)之间则差异显著，这说明遗传因素是药物代谢个体差异的决定因素。

药物摄入机体后，经过吸收、分布、与细胞(受体)相互作用，产生效应，经过生物转化、排泄而被机体消除。这一过程的各个环节都与一定的酶、受体或蛋白质的作用相关。例如，药物的吸收要借助于膜蛋白的转运，药物的分布要借助于血清蛋白的运输，药物作用于靶细胞要通过受体，药物的生物转化要经过酶促反应等。而这些与药物代谢、药物反应有关的酶、受体和蛋白质的合成是受基因所控制的，一旦这些基因出现异常，就会导致相应的药物代谢、药物反应异常。

遗传因素对药物代谢、药物反应的控制分为2类：一类受单基因控制；另一类受多基因控制。受多基因控制的药物代谢、药物反应研究较少；受单基因控制的药物代谢、药物反应研究较多，遗传机制也较清楚，本章重点介绍。

一、异烟肼代谢

异烟肼(isoniazid)是异烟酸和肼的化合物。该药对结核杆菌有强大的选择性抑制和杀灭作用，口服易吸收，疗效好，适应证广，是临床上常用的抗结核药。异烟肼在体内主要

通过 N—乙酰基转移酶(N - acetyltransferase)催化生成乙酰化异烟肼而失活(图 12 - 1),然后由肾排出。

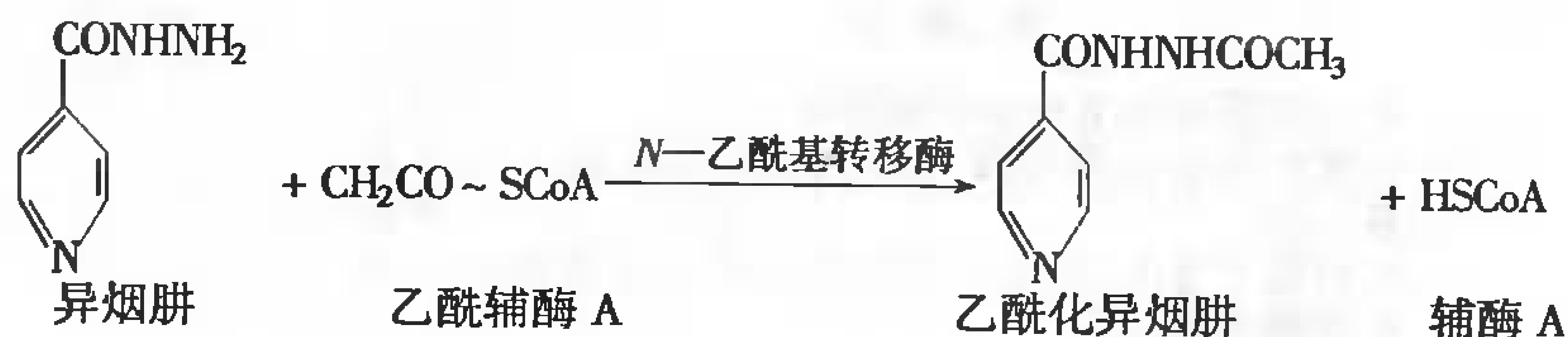


图 12 - 1 异烟肼的乙酰化

不同个体对异烟肼代谢的速度差异很大,口服异烟肼后,药物灭活较快,血浆半衰期仅为 45 ~ 80 min 者称为快灭活者(rapid inactivator);口服异烟肼后,药物灭活缓慢,血浆半衰期长达 2 ~ 4.5 h 者称为慢灭活者(slow inactivator)。慢灭活是由于 N—乙酰基转移酶的遗传性缺陷所致。家系调查表明,该酶受常染色体显性基因控制,快灭活者是显性纯合子,慢灭活者是隐性纯合子,杂合子则具有中等乙酰化速度。

N—乙酰基转移酶基因簇位于 8pter - q11,共有 *NAT*1、*NAT*2 和 *ANTP* 3 个基因,其中 *ANTP* 是假基因,*NAT*1 和 *NAT*2 有编码功能且高度同源(87%)。*NAT*1 无遗传变异性,编码的酶催化某些芳基胺药物的乙酰化;*NAT*2 有遗传多态性,编码的酶催化异烟肼等药物的乙酰化。*NAT*2 基因的 3 种突变型(M_1、M_2、M_3)可导致 N—乙酰基转移酶不稳定,活性降低,占慢灭活型的 95%以上。

异烟肼乙酰化速度既影响该药的疗效,也影响该药的不良反应。在疗效方面,由于快灭活者异烟肼血浆浓度较低,所以痰菌消失慢,且易出现耐异烟肼菌株。乙酰化速度对结核病疗效的影响很大程度上取决于治疗方案:每日给药 1 次,对疗效影响不大;每周给药 1 次,快灭活者较难治愈,且复发率较高。相比之下,慢灭活者疗效较好。在不良反应方面,若长期服用异烟肼,80%的慢灭活者发生多发性神经炎,而快灭活者只有 20%发生多发性神经炎。这是由于异烟肼在体内与维生素 B_6 反应生成异烟腙,维生素 B_6 失活,导致维生素 B_6 缺乏性神经损害所致。慢灭活者乙酰化速度慢,所以比快灭活者更易出现药物累积而导致多发性神经炎。同时服用异烟肼和维生素 B_6 可预防这种不良反应。此外,长期服用异烟肼还可导致部分患者发生肝炎,甚至肝坏死。在异烟肼引起的肝炎中,86%是快灭活者,这是由于快灭活者乙酰化速度快,可迅速产生大量的乙酰化异烟肼;乙酰化异烟肼在肝被水解为异烟酸和乙酰肼,乙酰肼在肝内又形成乙酰作用物质,导致肝组织坏死。

在不同地区、不同种族,异烟肼慢灭活者比例差异较大。埃及人高达 83%,美国人为 50% ~ 55%,黄种人为 5% ~ 20%,爱斯基摩人最少,仅为 5%左右。

除异烟肼外,N—乙酰基转移酶乙酰化作用灭活的其他药物,也可根据不同的灭活速度分为快速代谢和慢速代谢 2 种类型,而不同的代谢类型也具有不同的临床效应(表 12 - 1)。

表 12-1　乙酰化多态性与药物的临床效应

药物	临床效应	
	慢乙酰化	快乙酰化
异烟肼	常规剂量易发生外周神经性疾病 与苯妥英钠合用易发生苯妥英钠引起的不良反应 与利福平合用，西方人易发生血浆胆红素和转氨酶升高	中国人和日本人肝毒性作用甚为普遍 开放性结核 1 周 1 次治疗方案疗效较差
肼曲嗪（肼苯哒嗪）	可产生抗核抗体和系统性红斑狼疮样综合征	治疗高血压需较大剂量
柳氮磺胺吡啶	治疗风湿性关节炎疗效较好 不良反应，特别是血液方面和胃肠道不良反应发生率增高	高铁血红蛋白浓度升高
氨苯砜	血液方面不良反应多	治疗疱疹样皮炎需较大剂量
普鲁卡因酰胺	易发生系统性红斑狼疮样综合征，或发生较早	常规剂量治疗心脏病患者时，易发生室性期前收缩

二、琥珀酰胆碱敏感性

琥珀酰胆碱(succinylcholine)是一种麻醉辅助药，有肌松作用，能在神经肌接头处阻断神经冲动传递到骨骼肌纤维，使肌张力下降，骨骼肌松弛，便于进行外科手术。

琥珀酰胆碱的作用时间很短，很快被血浆及肝中的伪胆碱酯酶水解而失效。琥珀酰胆碱先被水解成无药理作用的琥珀酰单胆碱(图 12-2)，后者再被水解成琥珀酸和胆碱。不经过转化直接经尿排出的原药一般少于 2%。

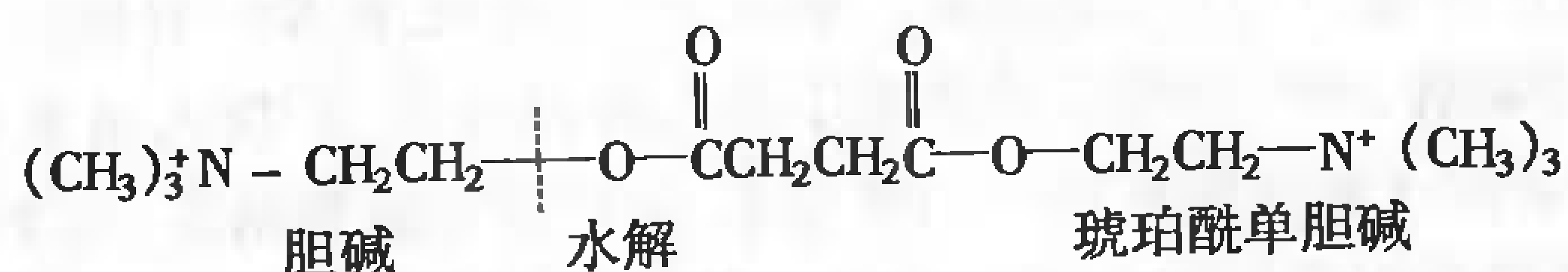


图 12-2　琥珀酰胆碱的降解

不同个体对琥珀酰胆碱的反应并不一致。通常静脉滴注常规药量使呼吸肌麻痹、呼吸暂停仅持续 2~3 min，随即恢复正常。极少数人(约 1/2 000)接受常规药量后，呼吸停止时间长达 1 h 或几小时，严重者甚至引起死亡，这样的个体为琥珀酰胆碱敏感性(succinylcholine sensitivity)个体。在这种情况下，可以通过输注同型正常人血浆或注射伪胆碱酯酶，恢复其呼吸功能。研究表明，这种异常反应是由于患者血中伪胆碱酯酶活性缺乏，不能以正常速度水解琥珀酰胆碱，导致神经肌接头处原药过多，引起呼吸肌持续性麻痹所致。琥珀酰胆碱敏感性属常染色体隐性遗传，这种敏感性个体在其他方面并无异常。

血浆伪胆碱酯酶(pseudocholinesterase)又称丁酰胆碱酯酶(butyrylcholinesterase，BCHE)，是由 4 个相同的亚基组成的四聚体。*BCHE* 基因定位于 3q26.1~q26.2，全长约 80 kb，含 4 个外显子，cDNA 长 2.4 kb。*BCHE* 基因突变类型很多，并且常有一个个体该基因上多位点突变现象，不同类型突变的发生率因人种而异。最常见的非典型突变是第 2 外显子第

209 位 A→G 突变，该突变使第 70 位天冬氨酸被甘氨酸取代，引起酶活性轻度降低。约 3% ~ 4% 的白种人为该型突变的杂合子，东方人和黑种人比较少见。

基因突变产生的变异型伪胆碱酯酶与其底物的结合力比正常伪胆碱酯酶低，因而水解琥珀酰胆碱的效力差。另外，许多变异型伪胆碱酯酶对伪胆碱酯酶的抑制剂如二丁卡因(dibucaine)、氟化物(fluoride)有较大的抗性，可通过用苯甲酰胆碱做底物，地布卡因和氟化物做抑制剂，来检测伪胆碱酯酶的活性。

三、葡萄糖－6－磷酸脱氢酶缺乏症

葡萄糖－6－磷酸脱氢酶(glucose－6－phosphate dehydrogenase，G6PD)缺乏症是一种常见于热带、亚热带地区的酶缺陷性遗传病。该病呈世界性分布，但各地区、各民族间的发病率差异很大，高发地区为地中海沿岸国家、印度东部、菲律宾、巴西和古巴等。在我国，该病主要见于长江流域及其以南各省，广东、广西、云南、四川、福建等省发病率较高。其中广东汉族人发病率高达 8.6%。G6PD 缺乏症以溶血性贫血为主要临床表现。患者平时无症状，但在食用蚕豆或服用伯氨喹啉类药物后出现血红蛋白尿、黄疸、贫血等急性溶血的临床表现，溶血具有自限性，一般都能恢复，但偶尔可引起致命后果。

G6PD 是细胞内葡萄糖分解代谢的次要途径——磷酸戊糖旁路的关键酶(图 12－3)。葡萄糖经过磷酸戊糖旁路代谢可生成磷酸核糖、还原型辅酶Ⅱ(NADPH)和 CO_2。其中 NADPH 作为供氢体参与体内多种代谢反应，包括维持谷胱甘肽(glutathione)的还原状态。还原型谷胱甘肽(GSH)是体内重要的抗氧化剂，可将体内的氧化剂(如 H_2O_2)还原，从而使一些含巯基(—SH)的蛋白质或酶免受氧化剂尤其是过氧化物的损害，同时其自身被氧化成氧化型谷胱甘肽(GSSG)。GSSG 在 NADPH 供氢的条件下，又可被还原成 GSH，再次发挥抗氧化作用。对于红细胞来说，GSH 的作用更为重要，它可以保护血红蛋白和红细胞膜蛋白的完整性。

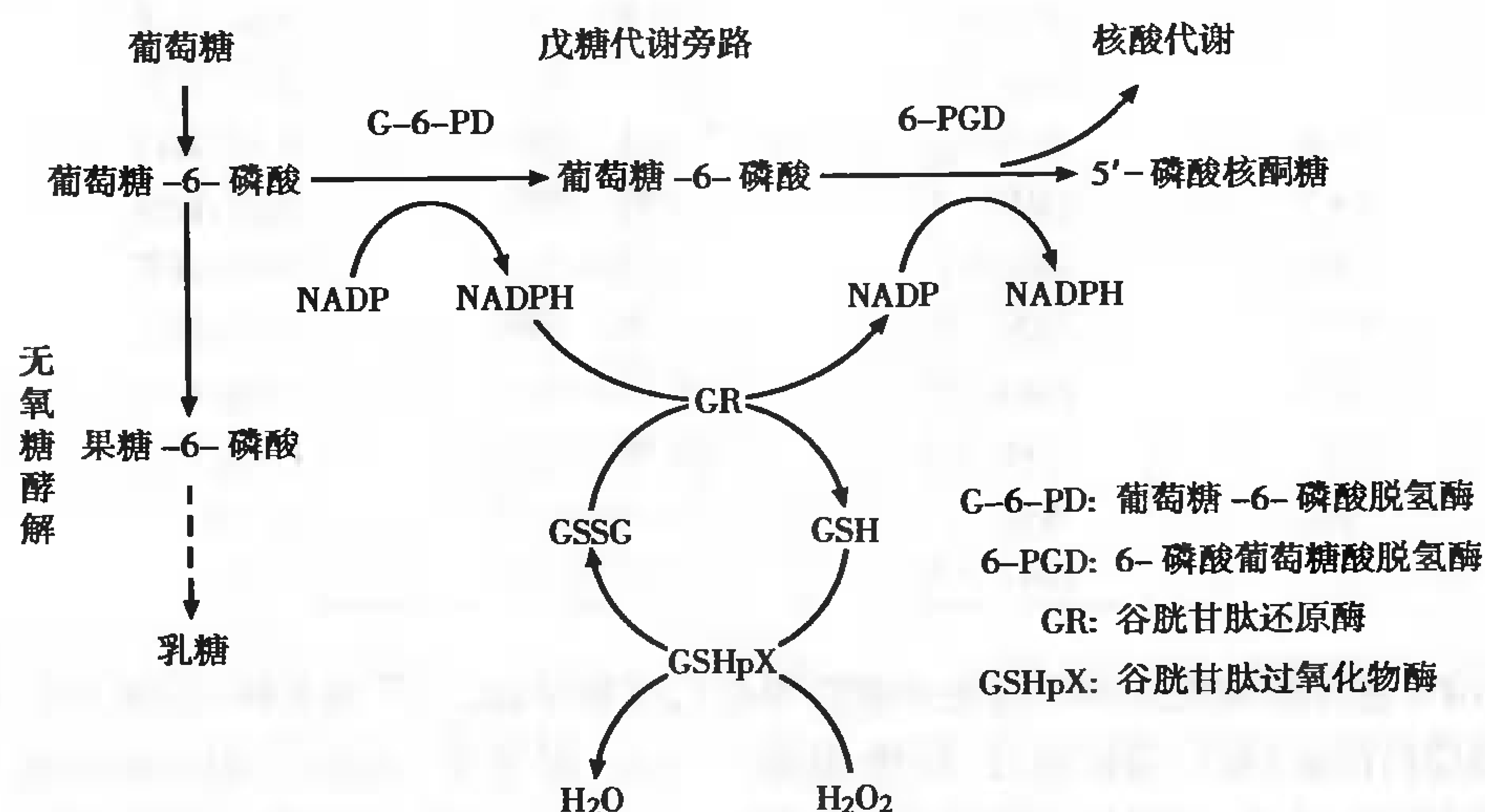


图 12－3 红细胞的糖代谢途径

红细胞中的糖代谢主要是通过无氧酵解进行，但也有 10%通过磷酸戊糖旁路代谢。这就为红细胞提供了足够的供氢体——NADPH。当 G6PD 缺乏时，磷酸戊糖旁路不能正常进行，NADPH 生成减少，GSH 含量降低，H_2O_2 可迅速将 GSH 耗尽，以至细胞内大量的 H_2O_2 不能被还原。这些过多的 H_2O_2 首先使血红蛋白 β 链表面第 93 位氨基酸（半胱氨酸）的巯基氧化，随即血红蛋白的 4 条多肽链因接触面不稳定而散开，H_2O_2 将进一步氧化血红蛋白内部的巯基，使血红蛋白变性，形成变性珠蛋白小体（Heinz 小体），并附着于红细胞膜上。此外，H_2O_2 还可氧化红细胞膜上的巯基，故这种红细胞容易被破坏。最近研究表明，NADPH 的减少本身也降低了红细胞对 H_2O_2 的抵抗性。由于上述原因，红细胞变形性降低，不易通过脾窦或肝窦而遭阻留破坏，引起血管内、外的溶血。

G6PD 基因定位于 Xq28，含 13 个外显子，全长 18 kb，编码 531 个氨基酸。其突变绝大多数为点突变，其中 50%以上为 G→C 颠换。G6PD 缺乏症属 X 连锁不完全显性遗传，男性半合子酶显著缺乏，女性杂合子酶活性变异范围大，可接近正常也可显著缺乏。

G6PD 基因突变所产生的 G6PD 生化变异型已达 400 多种，已鉴定的 *G6PD* 基因突变型 98 种，中国人中已发现 11 种突变型（表 12－2）。根据酶活性及临床表现，将 G6PD 缺乏症分为 5 类：①酶活性严重缺乏伴有代偿性慢性溶血，此类属非球形红细胞溶血性贫血，酶活性几乎为 0，无诱因亦可发生慢性溶血；②酶活性（ < 10%）严重缺乏，患者食用蚕豆或服用伯氨喹啉类药物后可诱发溶血；③酶活性（10% ~ 60%）轻度至中度缺乏，患者服用伯氨喹啉类药物可致溶血；④酶活性（60% ~ 100%）轻度降低或正常，此类患者一般不发生溶血；⑤酶活性增高，此类罕见，无临床表现。

表 12－2　中国人中的 *G6PD* 基因突变型

类型	碱基替换	氨基酸改变	发生的地区
C1	1376G→T	459 精→亮	大陆、台湾
C2	1388G→A	463 精→组	大陆、台湾
C3	1311C→T	无	大陆、台湾
C4	392G→T	131 甘→缬	大陆、台湾
C5	1024C→T	341 亮→苯丙	大陆、台湾
C6	95A→G	32 精→组	大陆、台湾
C7	592C→T	198 精→半胱	大陆、台湾
CT1	835A→T	297 苏→丝	美籍华人
CT2	1360C→T	454 精→半胱	大陆、台湾
CT3	493A→G	165 冬胺→天冬	台湾
	1004C→A	335 丙→天冬	大陆

G6PD 遗传性缺乏是某些常见药物性溶血的发病基础，也是蚕豆病、某些新生儿黄疸、某些感染性溶血（如细菌性肺炎、伤寒、病毒性肝炎、腮腺炎、流感等）的发病基础。G6PD 缺乏症明确诊断者，应避免进食蚕豆及其制品，忌服有氧化作用的药物，防治各种感染以预防或减少溶血的发生。表 12－3 中列出了可诱发 G6PD 缺乏者发生溶血的药物。

表 12-3 可诱发 G6PD 缺乏者溶血的药物

类 别	药 物
氨基喹啉类	伯氨喹啉、氯喹、扑疟喹啉、戊氨喹啉
砜类	氨苯砜、亚磺氨苯砜、噻唑砜
磺胺类	氨苯磺胺、磺胺醋酰、磺胺异噁唑、柳氮磺胺吡啶、磺胺对甲氧嘧啶
硝基呋喃类	呋喃妥因、呋喃唑酮、呋喃西林
镇痛药	乙酰水杨酸、非那西丁、乙酰苯胺
其他	维生素 K(水溶性同类物)、萘、羧苯磺胺、二巯基丙醇、亚甲蓝、乙酰苯肼、苯肼、氨甲苯酸、萘啶酸、新砷凡那明、奎宁、奎尼丁、氯霉素

四、无过氧化氢酶血症

过氧化氢溶液是外科常用的清创药,它在组织中过氧化氢酶的作用下能迅速分解,释放出氧,起抗菌除臭的作用。正常情况下,用过氧化氢溶液消毒的创面呈鲜红色,有泡沫产生。1946 年,日本耳鼻喉科医生高原(Takahara)首次报告,在给一名患口腔坏疽的女孩用过氧化氢消毒创面时,发现创面变成棕黑色,且无泡沫产生,当时称之为“黑血病”。后来证实,患者红细胞中缺乏过氧化氢酶,不能将过氧化氢分解成水和氧气,而过氧化氢使血红蛋白氧化变性,生成棕黑色的高铁血红蛋白,故称本病为“无过氧化氢酶血症”(acatalasia)。约半数患者易患牙龈溃疡、牙龈坏疽、牙龈萎缩,甚至牙脱落。

家系分析表明,本病为常染色体隐性遗传病,但根据酶活性的差异,本病又表现出不完全显性的特征。患者(cc)酶活性丧失或活性很低,杂合子(Cc)酶活性介于正常人(CC)和患者之间。因此,也有人称本病为“过氧化氢酶缺乏症”。

过氧化氢酶基因定位于 11p13.5~p13.6,与 WAGR 综合征基因、钙调素基因和甲状腺素基因紧密连锁。根据酶的活性、基因频率、地理分布和临床表现,目前已发现 5 种无过氧化氢酶血症变异型。

本病黄种人发病率较高,日本某些地区发病率高达 1%,欧洲大陆也有发现(瑞士为 0.7%)。我国华北地区发病率为 0.65%,华中为 0.55%,华南为 0.23%,台湾为 0.29%。

五、异喹胍-金雀花碱代谢多态性

降压药异喹胍(debrisoquine)和具有抗心律失常及催产作用的药金雀花碱(sparteine)在不同个体的代谢率差异很大,它们都由细胞色素氧化酶 P4502D6(CYP2D6)氧化代谢而失活。

异喹胍在体内经 CYP2D6 代谢,生成 4-羟基异喹胍,不同个体的 CYP2D6 代谢能力不同,故有强代谢者(extensive metabolizer, EM)和弱代谢者(poor metabolizer, PM)之分。测定 CYP2D6 代谢能力的方法是:晚上口服异喹胍后收集 8 h 的尿,检测尿中异喹胍原型和 4-羟基异喹胍的量,二者相比即为异喹胍代谢率(MR)。MR 大于 12.6 者为 PM,小于 12.6者为 EM。

$$MR = \frac{\text{尿中异哇胍原型的量}}{\text{尿中 4-羟基异哇胍的量}}$$

CYP2D6 的弱代谢表型(PM)为常染色体隐性遗传。*CYP2D6* 基因定位于 22q13.1 ~ qter,含 9 个外显子,全长 7 kb。当 *CYP2D6* 基因发生 *A*、*B*、*D* 等位基因突变时,可导致野生型基因(*CYP2D6* – wt)活性消失,PM 表型出现。2*D6* – *A* 突变是第 5 外显子第 2637 位 A 丢失,导致移码突变;2*D6* – *B* 突变是第 3 内含子 3′末端第 1934 位 G→A 转换,导致内含子外显子接头处剪切位点改变。这两类突变可提前终止翻译而导致酶蛋白的末端合成障碍。2*D6* – *D* 突变是指 *CYP2D6* 基因完全缺失。异喹胍 PM 的突变基因大多是 2*D6* – *B*(占 75%)和 2*D6* – *A*(占 5%)。此外,一些突变率较低的弱代谢突变基因,如 *CYP2D6E*、*CYP2D6* * 6、*CYP2D6T* 等也可导致 *CYP2D6* – wt 活性丧失,PM 表型出现。

在不同遗传背景的人群中,PM 的比例差异显著,尼日尼亚人为 15%,加拿大人为 12%,英国人为 9%,瑞典人为 5.4%,埃及人为 1.5%,我国汉族人仅为 0 ~ 1%。因此,遗传决定的某些药物氧化代谢的多态性可能是不同种族患者所需相应药物剂量不同的重要原因。

由于药物代谢能力下降,可导致 CYP2D6 弱代谢者用药后毒副作用增多。而对于某些需要 CYP2D6 激活的药物,弱代谢者则不容易达到预期的药效。例如,弱代谢者使用常规剂量的可待因,通常起不到明显的止痛效果。

另外,据研究,CYP2D6 的多态性与某些疾病有相关性,弱代谢者易发生红斑狼疮和帕金森综合征;而强代谢者易发生肺癌、膀胱癌、肝癌和胃肠癌。其机制尚不清楚。

除异喹胍和金雀花碱由 CYP2D6 氧化代谢外,还有一些药物的代谢受 CYP2D6 控制,例如,普萘洛尔、奎尼丁、普罗帕酮、可待因、丙咪嗪等。

第二节　毒物反应的遗传基础

药物仅是人类接触的环境化学因子的一小部分,环境中还存在许多其他的潜在毒物和对某些具有特定遗传素质的人发生毒性损伤的物质。与药物一样,这些物质在人体内的代谢途径也受到特定基因型的制约,不同个体产生不同的效应。例如,同样吸烟,有的人患肺癌,有的人不患肺癌。研究群体中不同个体对各种环境因子的特殊反应方式和适应特点的遗传基础的学科称为生态遗传学(ecogenetics),是从药物遗传学发展而来的遗传学分支学科。

一、乙醇中毒

人们对乙醇的耐受存在种族和个体差异。乙醇敏感者,当摄入 0.3 ~ 0.5 ml/kg 乙醇时,就出现面红耳赤、皮温升高、心率加快等急性乙醇中毒表现;乙醇耐受者则无此表现。黄种人 80% 为乙醇敏感者,白种人仅 15% 为乙醇敏感者。

乙醇在体内的代谢主要包括 2 步反应:第 1 步是乙醇在肝乙醇脱氢酶(alcohol dehydrogenase,ADH)作用下生成乙醛,第 2 步是乙醛在乙醛脱氢酶(acetaldehyde dehydrogenase,ALDH)作用下进一步代谢为乙酸。两步反应式如下:



$$C_2H_5OH + NAD^+ \xrightarrow{ADH} CH_3CHO + NADH + H^+$$



$$CH_3CHO + H_2O + NAD^+ \xrightarrow{ALDH} CH_3COOH + NADH + H^+$$

第1步反应生成的乙醛对肾上腺素、去甲肾上腺素等物质的分泌有刺激作用，故急性酒精中毒患者有上述各种表现。

ADH由二聚体构成，现已发现至少有5种基因（ADH_1基因、ADH_2基因、ADH_3基因、ADH_4基因和ADH_5基因）编码人体乙醇脱氢酶8个多肽亚单位（α、β_1、β_2、β_3、γ_1、γ_2、π和χ），这些亚单位成对组成不同乙醇脱氢酶的二聚体。ADH_1基因、ADH_2基因和ADH_3基因位于4q21－q24，在不同的组织和不同发育时期差异表达，成人期主要是ADH_2基因的表达。ADH_2具有多态性，大多数白种人ADH_2基因编码的乙醇脱氢酶由$\beta_1\beta_1$组成，而90%黄种人的乙醇脱氢酶由$\beta_2\beta_2$组成，β_2是ADH_2基因突变产生的多肽亚单位，与β_1仅有1个氨基酸的差异（$\beta_1$47位精氨酸→$\beta_2$47位组氨酸）。$\beta_2\beta_2$酶活性约为$\beta_1\beta_1$的100倍，所以，大多数白种人在饮酒后产生乙醛较慢，而黄种人饮酒后乙醛蓄积速度较快。这是黄种人易出现乙醇中毒的原因之一。

乙醇代谢过程的另一关键酶——乙醛脱氢酶主要有2种同工酶：$ALDH_1$和$ALDH_2$，二者均为多肽四聚体。$ALDH_1$存在于胞质中，基因定位于9q21；$ALDH_2$存在于线粒体内，基因定位于12q24.2。$ALDH_2$具有遗传多态性，其缺乏者属常染色体显性遗传。$ALDH_2$基因长44 kb，含13个外显子，编码517个氨基酸。基因突变发生在第12外显子，导致相应多肽链第487位谷氨酸被赖氨酸取代，形成无功能的$ALDH_2$，不能有效代谢乙醛。$ALDH_2$活性比$ALDH_1$高，白种人几乎无$ALDH_2$功能缺失，而50%的黄种人$ALDH_2$功能缺失，氧化乙醛的速度慢，这是黄种人易出现乙醇中毒的又一原因。

综上所述，黄种人较白种人易产生乙醇中毒的原因是由遗传因素决定的。具有ADH_2（$\beta_2\beta_2$）及$ALDH_1$者对乙醇最敏感；具有ADH_2（$\beta_1\beta_1$）及$ALDH_1$者次之；具有ADH_2（$\beta_1\beta_1$）及$ALDH_2$者最不敏感。大多数黄种人在饮酒后产生乙醛速度快而氧化为乙酸的速度慢，故易产生乙醛蓄积而导致乙醇中毒。大多数白种人则与此相反。

二、吸烟与肺癌

肺癌是最常见恶性肿瘤之一，吸烟被公认为是肺癌的重要危险因素。流行病学调查证明，80%～90%的男性肺癌与吸烟有关，但并不是所有吸烟者都患肺癌，吸烟者是否患肺癌与个体的遗传基础有关。

烟草中含有多种致癌物质，多环苯蒽化合物是其中之一。多环苯蒽化合物本身致癌性较弱，进入机体后，通过细胞内芳烃羟化酶（aryl hydrocarbon hydroxylase，AHH）的作用可转变为具有较高致癌活性的7，8—二羟基—9，10—环氧芘，促进肺癌的发生。同时，苯蒽化合物能提高AHH活性，是AHH的诱导剂。因此，AHH的可诱导性具有增强苯蒽衍生物致癌性的作用，受遗传因素决定，在人群中呈多态性分布，可区分为高、中、低3个诱导活性组。目前认为，AHH的可诱导性可能受2号染色体短臂一对等位基因控制，按孟德尔方式遗传。另有研究表明，如以低诱导活性组肺癌易感性为1，则中诱导活性组达16，而高诱导活性组则高达36。

第三节　药物基因组学

药物反应存在个体差异是是客观存在的,药物遗传学的主要任务就是要阐明遗传因素在机体对药物和外源性物质代谢和反应中的作用。近年来,新的分子生物学技术的应用以及人类基因组计划的顺利实施,使药物遗传学的研究和发展进入了一个新阶段。伴随着功能基因组学的提出和高通量筛选方法的出现,产生了一个新的学科——药物基因组学(pharmacogenomics)。药物基因组学以药物安全性为目标,研究各种基因突变与药效及安全性之间的关系,利用基因组学的知识,根据不同人群及不同个体的遗传特征来设计和制造药物,最终达到个体化治疗的目标。药物基因组学所要研究的是决定药物行为和敏感性的整个基因表达谱,而药物遗传学则主要研究药物作用谱中狭窄的药物代谢和反应中遗传差异的作用机制。但这种区分是人为的,这两个学科术语经常交换使用。

药物基因组学的研究目前正朝向应用生物技术和药物工业化方法,通过加速鉴定疾病相关靶点以促进新药的发现方向发展。在基因组上结合化学和高效扫描技术寻找这些靶点,可为发现和鉴定更多新化合物创造条件。

复杂遗传疾病的基因组研究包括连锁分析和关联分析。由于从多个家系中很难获得药物反应方面的资料,所以,连锁分析不适合用于药物基因组学研究。关联分析是在不相关的人群中寻找与性状(疾病或药物反应)相关的染色体区域。由于开放人群中不相关个体的共同祖先较家系成员亲缘关系更远,因此共有染色体区域很短,在 100 kb 或 100 kb 以下。在开放人群中进行基因组关联分析,至少需要 10 万个标记或更多,而随着 DNA 标记发现的速度越来越快及新的分型方法的应用,使在基因组范围内进行关联分析成为可能。目前,虽然用单核苷酸多态性(single nucleotide polymorphism, SNP)为遗传标记进行关联分析较微卫星标记有更多的优越性;但利用 SNP 进行的关联分析具有存在遗传异质性的缺点,并且也不适合用于具有稀少等位片段的基因的检测。

药物反应多态性与 RNA 和蛋白质的改变有关。RNA 的变化表现为不同个体、不同时期的转录水平的差异,使用基因芯片技术可进行大规模的转录分析。转录分析在癌症研究中侧重于表达分析,在其他疾病中可用于鉴定侯选基因。同时,它能在大量患者人群中对侯选基因多态性进行扫描,并能预测药物反应的基因表达情况,也可通过抗体对唾液和血清中的蛋白水平进行检测。蛋白质的变化一方面可由转录水平的变化来预测,另一方面可以通过 ELISA、双向 PAGE 凝胶电泳及质谱分析等蛋白质组学方法对临床样品进行大规模分析,从而鉴定出高预测性的药物基因组学蛋白标记。

药物基因组学下一阶段的任务包括:针对决定药物反应差异的编码蛋白基因建立药物遗传学多态性数据库;系统准确地鉴定一套频率大于 1% 的公用遗传多态性标记;在一套合适的人类材料上对公用多态性区域进行完整测序,包括外显子、调节区和内含子序列。同时,鼓励个人建立、维护样品库以便共享;建立包括 DNA、组织和血液样品的资源库;对患者应准确诊断对药物反应的表型,并收集病理结果、病历和父母及家庭资料;资助基础研究实验室建立检测 DNA 序列多态性技术,对正常人群和患者人群区分功能性显著差异的遗传多态性;在生物信息学、临床药理学、人类遗传学等方面加强人员培训等。

总之，现代医学不再只注重预防和治疗，正朝向安全有效和药物的经济治疗方向发展，即集中在研究药物作用的遗传分布来满足临床需要，而药物基因组学开辟了一个崭新的领域，在药物设计、制造和应用方面正酝酿着一场根本性的革命。它将目前依据患病人群共性的药物治疗转向根据不同人群及不同个体遗传特征来设计、制造药物进行治疗，从而使人类进入到一个崭新的个性化医疗的时代。

小　结

药物遗传学属于生化遗传学的范畴，其主要任务是阐明遗传因素在机体对药物和外源性物质代谢和反应中的作用。遗传因素对药物代谢、药物反应的控制可分为单基因控制和多基因控制两类，受单基因控制的药物代谢、药物反应研究较多，遗传机制也较清楚。

异烟肼代谢分为快灭活和慢灭活两类。慢灭活能力属常染色体隐性遗传，是位于8pter－q11 的 *NAT*2 基因点突变导致肝脏 *N*－乙酰基转移酶活性降低所致。琥珀酰胆碱敏感性是由位于 3q26.1－q26.2 的 *BCHE* 基因突变使血清胆碱酯酶亲和力降低所致。G6PD 缺乏症属 X 连锁不完全显性遗传。G6PD 基因定位于 Xq28，基因突变使 G6PD 缺乏，红细胞磷酸戊糖旁路代谢不能正常进行，红细胞抗氧化能力下降，易被破坏而发生溶血性贫血。无过氧化氢酶血症属常染色体隐性遗传，其基因突变使该酶缺乏而出现相应症状。异喹胍－金雀花碱等药物的代谢受控于 CYP2D6，可分为强代谢者(EM)和弱代谢者(PM)。*CYP*2*D*6 基因位于 22q13.1－qter，具有多态性。该位点的 2*D*6－*A*、2*D*6－*B*、2*D*6－*D* 等基因突变时可导致酶活性丧失，PM 表型出现。

除药物外，环境中其他的化学因子在人体内的代谢途径也受到特定基因型的制约，使不同个体产生不同的效应。例如，ADH 和 ALDH 的合成受遗传因素决定，具有 $ADH_2(\beta_2\beta_2)$ 及 $ALDH_1$ 者对乙醇最敏感；具有 $ADH_2(\beta_1\beta_1)$ 及 $ALDH_1$ 者次之；具有 $ADH_2(\beta_1\beta_1)$ 及 $ALDH_2$ 者最不敏感。大多数黄种人在饮酒后产生乙醛速度快而氧化为乙酸的速度慢，故易产生乙醛蓄积而导致乙醇中毒。再如，不同吸烟个体患肺癌的风险与 *AHH* 基因控制的 AHH 可诱导性的大小有关。

药物基因组学是在人类基因组计划顺利实施的背景下，药物遗传学所进入的一个崭新领域。它以药物安全性应用为目标，研究各种基因突变与药效及安全性的关系，利用基因组学的知识，根据不同人群及不同个体的遗传特征来设计和制造药物，最终达到个体化治疗的目标。

(余　红)

第十三章　人类基因的研究技术

分子生物学的迅猛发展，特别是基因研究技术的日臻完善，极大地推动了医学遗传学的发展。重组 DNA 技术、PCR 技术、核酸分子杂交技术以及基因定位技术的应用，不仅使人们对遗传病致病基因的定位与克隆、分子机制的认识不断深入，而且使遗传病的基因诊断和基因治疗也取得了突破性进展。

第一节　重组 DNA 技术——基因工程

重组 DNA 技术(recombinant DNA technique)是现代分子生物学技术发展的重要成就之一，也是基因工程(gene engineering)的核心技术。基因工程是指人类根据自己的需要选择目的基因(DNA 片段)，在体外对其进行剪切和重新选择，构成重组 DNA 分子，然后把它导入宿主细胞，从而扩增有关 DNA 片段，表达其基因产物，制备特定蛋白，或是特定改造生物个体特性。

一、工具酶

在基因工程研究中，用以切割、连接和修饰 DNA 或 RNA 的一类酶称为工具酶。包括：限制性内切核酸酶、DNA 连接酶、末端转移酶、逆转录酶等(表 13－1)，其中最主要的是限制性内切核酸酶。

表 13－1　重组 DNA 技术中最常用的工具酶

工具酶	主要用途
限制性核酸内切酶	识别 DNA 特定序列，切断 DNA 链
DNA 聚合酶 Ⅰ	缺口平移，制作标记 DNA 探针
DNA 聚合酶 Ⅰ 大片段(Klenow)	①补齐双链 DNA 的 3′末端 ②通过补齐双链 DNA 的 3′末端使 3′末端标记 ③在 cDNA 克隆中，第二链的合成 ④DNA 序列分析
T4DNA 连接酶	连接 3 个 DNA 分子或片段
逆转录酶	构建 cDNA 文库
碱性磷酸酶	防止载体自身连接，提高重组效率
末端转移酶	在 3′末端加入同质多聚物尾，以便于连接
耐热 DNA 聚合酶(Taq DNA 聚合酶等)	聚合酶链反应(PCR)
多核苷酸激酶	催化多核苷酸 5′羟基末端磷酸化，制备末端标记探针

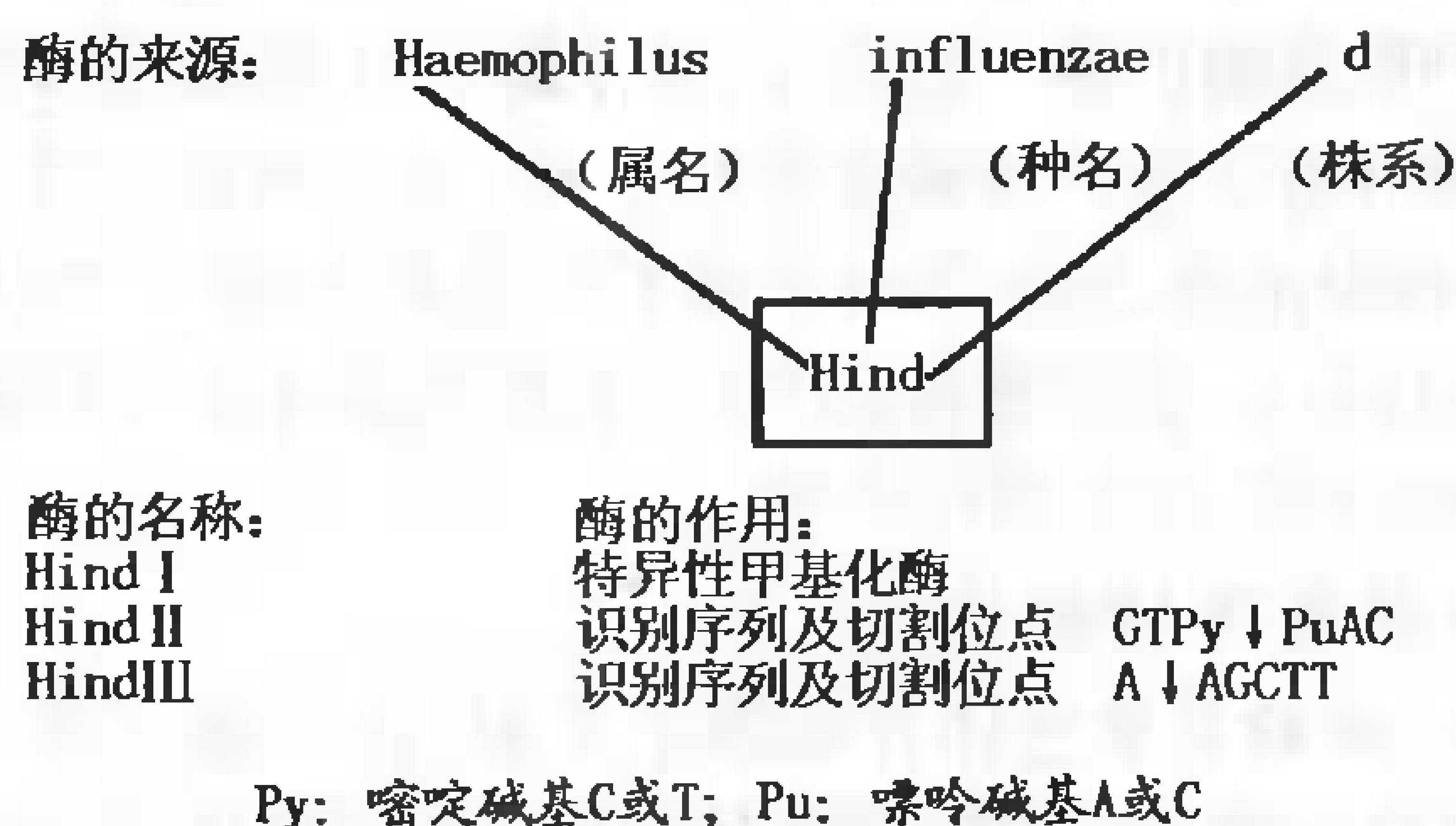
(一)限制性内切核酸酶

能识别DNA特异序列,并在识别位点或其周围切割DNA双链结构的一类内切酶,称为限制性内切核酸酶(restriction endonuclease),简称限制酶。自1970年Smith首次证明限制酶hindⅡ及它的特异性切割位点至今,已纯化限制酶1 000多种,已证实有特异性切割位点的有150多种。迄今为止发现的限制酶均来源于原核生物,几乎所有的原核生物都含有这类酶。

1.限制酶的分类 根据酶的结构、作用及与DNA结合和切割位点的特异性,限制酶被分为3型,其中Ⅰ型和Ⅲ型因其识别和切割位点不在同一部位,而且切割产物无特异性,所以在基因工程中不常用;由于Ⅱ型限制酶能在DNA分子内部的特异位点识别和切割双链DNA,切割的产物也是特定的。所以,Ⅱ型限制酶成为基因工程中最重要的工具酶。通常基因工程中所用的限制酶,即指这类酶。

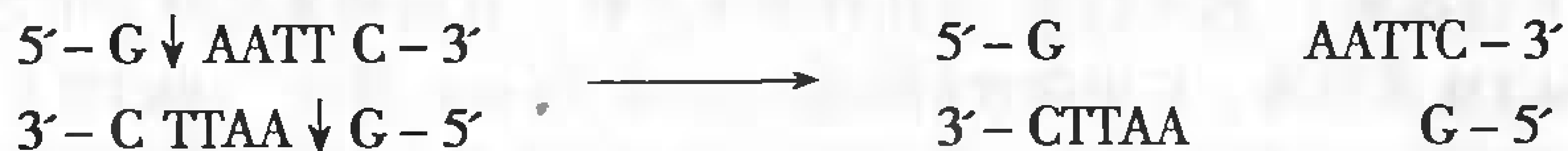
2.限制酶的命名 限制酶的命名遵照如下规则:①按酶来源的相应微生物学名,取其属名的第1个大写字母和种名的第1、2两个字母(小写)组成酶的基本名称;②若微生物有不同的株系,取其株系的第1个字母加于酶的基本名称之后;③用大写罗马数字Ⅰ、Ⅱ、Ⅲ等来区分存在于同种微生物中具有不同特异性的限制酶。

例如,来源于流感嗜血杆菌的3种限制酶的命名:



Py: 嘧啶碱基C或T; Pu: 嘌呤碱基A或C

3.限制酶的识别和切割位点 绝大多数限制酶的识别位点通常是具有回文序列的DNA片段。多数限制酶是错位切割双连DNA,产生黏性末端(sticky end)。例如,EcoRⅠ切割后产生5′黏性末端:



PstⅠ切割后产生3′黏性末端:



另外一些酶可以沿对称轴切割DNA,产生平末端(blunt end)。例如,SmaⅠ切割后产生平末端:



(二)修饰酶

修饰酶(modifying enzymes)是指一类对 DNA 和 RNA 分子末端进行剪接、补平、连接以及化学修饰的酶,被广泛应用于各种基因操作的实验中。包括:DNA 连接酶、DNA 聚合酶、逆转录酶、碱性磷酸酶、末端脱氧核苷酸转移酶等多种。

二、载体

载体(vector)亦称运载体,是将外源目的 DNA 导入受体细胞,并能自我复制和增殖的工具。载体实际上也是 DNA,通过体外重组的方式,可以把目的基因和载体相连,构成 DNA 重组体,然后转移到适当的受体细胞中,进行无性繁殖,从而获得大量的基因片段或相应的蛋白质产物。没有合适的载体,外源基因很难进入受体细胞中,即使能够进入,也不能进行复制和表达。

目前,基因工程中所用的载体都是通过改造天然质粒、噬菌体和病毒等构建的。

(一)载体的特性

在基因工程中的载体主要有 5 类:①质粒,主要指人工构建的质粒;②λ 噬菌体的衍生物;③黏粒;④单链 DNA 噬菌体 M13;⑤动物病毒。

尽管各类载体的来源不同,在大小、结构及复制等方面的特性差别很大,但作为基因工程的载体,都应具备以下基本条件:①必须含有复制起始点,可在适当的宿主细胞中自主复制;②应具有多种限制酶的单一切点,以便外源 DNA 能通过共价键连接插入;③应具有筛选标记,以区别阳性重组分子和阴性重组分子。选择标记包括:抗药性基因、酶基因、营养缺陷型及形成噬菌斑的能力等;④分子量不宜过大,以便于 DNA 体外操作,同时载体 DNA 与宿主核酸应容易分开,便于提纯;⑤对于表达型载体还应配备相适应的启动子、前导序列、加尾信号、增强子等 DNA 调控元件。

(二)基因工程中常用的载体

1.质粒(plasmid) 质粒是染色体外能独立复制、稳定遗传的一种环状双链 DNA 分子。质粒 DNA 分子可以持续稳定地以游离状态存在于染色体外,但在特定的条件下又能可逆地整合到宿主的染色体上,随着宿主染色体的复制而复制,并通过细胞分裂传递到子1代。

质粒只能容纳小于 10 kb 的外源 DNA 片段,主要用作亚克隆载体。一般来说,外源 DNA 片段越长越难插入,越不稳定,转化效率也越低。常用的质粒有 pBR322、pUC 系列、pSP 系列和 pGEM 系列等。其中最常用的是 pBR322 和 pUC 系列。pBR322 长 4.3 kb,含有氨苄青霉素和四环素的抗性基因,在氨苄青霉素和四环素的抗性基因中间有限制酶酶切位点,以便于外源基因的插入和筛选。

2.噬菌体载体(phage vector) 噬菌体是一类细菌病毒的总称,能高效感染细菌,并在其中克隆中等长度的 DNA 片段,用作克隆载体的噬菌体主要有 λ 噬菌体、M13 噬菌体以及由 λDNA 的 cos 区与质粒重新构建的黏粒(cosmid)。

3.大片段 DNA 克隆载体 因为人和哺乳动物的基因组很大,无法插入质粒和噬菌体类载体,因此,在构建重组 DNA 分子时就需要能插入大片段 DNA 的载体。目前,这类载体主要有酵母人工染色体和细菌人工染色体。

(1)酵母人工染色体　人类基因组有 3×10^9 个碱基对,人类染色体的平均大小为 130 Mb,难以直接进行核苷酸序列分析。用黏粒和 λ 噬菌体克隆排序亦十分费力,因为需要相当多的黏粒克隆才能包含整个人类基因组。为解决这一难题,1987 年,Olson 实验室构建成酵母人工染色体(yeast artificial chromosome,YAC)克隆载体。它包括若干酵母基因片段和 pBR322 质粒的 DNA 片段,可克隆 1Mb 以上的 DNA 片段。YAC 在人类基因组研究中起了非常重要的作用,但随着研究工作的不断深入,YAC 也表现出一些缺点,如克隆效率低、易形成克隆嵌合体、插入片段不稳定以及难以制备纯的 YAC－DNA 等。

(2)细菌人工染色体　为克服 YAC 的不足,人们又构建了几种以大肠杆菌为宿主的替代文库,包括:细菌人工染色体(bacterial artificial chromosome,BAC)库、P1 噬菌体库和 P1 衍生人工染色体库。BAC 是 F 因子建立起来的单拷贝大肠杆菌载体。对大肠杆菌进行适当的突变,使 BAC 载体中插入的外源片段达到 300kb 也能稳定复制。BAC 载体的稳定性很好,它可用于克隆大片段 DNA,从而构建整个人类基因组文库,并且可快速而高精确度地分析插入的片段。

4.真核病毒载体　以上介绍的载体,除酵母载体外,均只适用于在细菌中繁殖和生存,远远满足不了真核 DNA 的基因工程技术的需要,特别是真核基因表达的需要。对真核基因而言,动物病毒载体是一种比较理想的体系。目前常用的动物病毒载体有:猴空泡病毒、牛痘病毒、腺病毒和逆转录病毒。

三、重组 DNA 技术的基本过程

重组 DNA 技术是基因工程的核心技术,可用于各种不同的目的,如:克隆某个特定的基因,建立基因文库,将特定的基因片段亚克隆以进行 DNA 测序,通过基因工程表达某已知的基因以获得其产物等。尽管重组 DNA 技术在用于不同目的时,其制备基因、选择载体以及鉴定重组体等方面有所不同,但基本过程均包含如下几个步骤。

(一)目的基因的制备

目的基因是指所要研究或应用的基因,也就是需要克隆或表达的基因。根据 DNA 重组体的不同目的,可采用不同的方法制备目的基因。

1.制备基因组 DNA　首先要选择含有目的基因的实验材料。如果实验目的是为了研究目的基因表达调控有关顺序(如内含子),那么应选用染色体 DNA 作为克隆基因的实验材料。采用合适的限制酶将基因组剪切成一定的片段,每一个 DNA 片段都与一个载体分子拼接成重组 DNA。将所有的重组 DNA 分子都引入宿主细胞并进行扩增,得到全部基因组中不同片段的分子克隆混合体系,这样一个混合体被称为基因文库(genomic library),然后通过杂交筛选从基因文库中获得特定的基因片段。

2.制备 cDNA　目的基因的 DNA 来源也可以选择由 mRNA 反转录得到的 DNA,即互补 DNA(complementary DNA, cDNA)。一般而言,如果实验目的是研究基因编码区的氨基酸顺序或者是获得能在原核系统中表达的真核基因,则应该选用 cDNA 作为克隆基因的材料。可以将 cDNA 的混合体与载体进行连接,使每一个 cDNA 分子都与一个载体分子拼接,形成 DNA 重组体。将这些重组 DNA 分子引入宿主细胞并进行扩增,得到分子克隆的混合体,即 cDNA 文库(cDNA library)。通过杂交筛选获得特定的 cDNA 克隆。

3. 制备用于表达的基因片段　随着分子生物学实验技术的飞速发展，人们对基因的认识也在不断的加深，基因的克隆正以每年上千条的速度增加。如果需要重新克隆这些已知核苷酸顺序的基因时，可以选择以下方法：①直接用限制酶切取；②用 PCR 技术体外扩增有关的 DNA 片段；③化学合成。

(二)载体的选择和制备

选择载体主要是根据构建重组体的目的，同时要考虑载体中应有合适的限制酶切位点。如果构建目的需要表达一个特定的基因，则要选择合适的表达载体。无论选择何种载体，首先要获得载体分子，即分离纯化 λ 噬菌体 DNA、制备黏粒或其他有关质粒。获得载体后，采用适当的限制酶将载体 DNA 进行切割，获得线性分子，以便于同目的基因片段进行连接。

(三)DNA 分子的体外连接

不同来源的 DNA 片段通过限制酶或机械力剪切后，可以在体外重新连接形成人工重组体。体外连接的方法有以下几种。

1. 黏性末端连接　大多数限制酶能将 DNA 分子切割，形成黏性末端。如果目的基因和载体用同一种限制酶剪切，则能产生相同的黏性末端，因而，DNA 片段之间很容易按碱基配对关系退火，在 T_4DNA 连接酶的作用下形成重组体。

2. 平末端连接　有些限制酶，如 Alu Ⅰ、Sam Ⅰ 等，能将 DNA 分子沿对称轴剪切，形成平末端片段。这些平末端的 DNA 片段在一定的反应条件下，通过 T_4 DNA 连接酶的作用，将其有效地连接起来。

3. 人工接头连接　DNA 平末端连接虽然适用范围很广，但是它存在效率低、非重组背景高、多拷贝插入以及双向插入等缺陷。这种情况下采用人工合成的接头进行连接比较有效。所谓接头是指含有某些限制酶切点的寡核苷酸片段，在 T_4 DNA 连接酶的作用下，将接头连接到目的 DNA 片段的两端，然后再用相应的限制酶切割，使外源 DNA 片段的两端形成黏性末端，这样比较容易同载体连接。

4. 同聚物加尾连接　当载体和目的基因没有可互补的黏性末端时，运用同聚物加尾连接的方法，可在载体和目的基因片段上造成互补的末端序列，再进行连接。这是连接 DNA 片段的一种有效的方法，被广泛应用于基因克隆研究中。

(四)外源 DNA 导入宿主细胞

根据重组 DNA 分子的大小，用不同基因转移方法，将它们导入宿主细胞，使重组 DNA 分子在宿主细胞内进一步扩增。常用的基因转移方法有以下几种。

1. 转化(transformation)　是指将质粒或其他外源 DNA 导入处于感受态的宿主细胞，并使其获得新的表型的过程。常用的宿主细胞为大肠杆菌。

2. 感染(infection)　是指以 λ 噬菌体、黏粒或真核细胞病毒为载体的重组 DNA 分子，在体外经过装配成为具有感染能力的病毒或噬菌体颗粒，然后感染适当的细胞，并在细胞内扩增的过程。由噬菌体和细胞病毒介导基因转移过程也称为转导(transduction)。

3. 转染(transfection)　是指真核细胞主动摄取或被动导入外源 DNA 片段而获得新的表型的过程。进入细胞的 DNA 可以被整合到宿主细胞的基因组中，也可以在染色体外存在和表达。用这种方法可以从基因组中筛选出具有某种功能的基因。

(五)目的基因的筛选和鉴定

将外源基因导入宿主细胞以后,首要任务就是筛选含有目的基因的转化菌,然后进行扩增。其步骤是:首先筛选出转化菌;然后筛选出带有重组体的克隆;最后对 DNA 重组体进行鉴定。所用的方法主要有:遗传学方法、免疫学方法、核酸杂交以及 PCR 技术等。

第二节 分子杂交

分子杂交(molecular hybridization)是指 2 条含有互补序列的核酸单链,在一定条件下,按碱基配对原则形成双链的过程。杂交的双方分别称为探针和待测核酸,杂交后形成的异源双链分子称为杂交分子。

分子杂交技术是目前生命科学领域中应用最广泛的技术之一,具有灵敏度高、特异性强等优点。主要用于特异 DNA 或 RNA 的定性或定量检测。如:测定特异 DNA 序列的拷贝数、特定 DNA 区域的限制性内切酶图谱,以判断是否存在基因缺失、插入等重排现象;末端标记的寡核苷酸探针可检测特定位点的基因突变;分子杂交还可以进行 RNA 结构的粗略分析、特异 RNA 的定量检测、特异基因克隆的筛选等。在分子克隆、基因诊断以及核酸序列分析等方面发挥着重要作用。

一、分子杂交的基本原理

分子杂交的基本原理是:具有互补序列的 2 条单链核酸分子在一定条件下(适量的温度及离子强度),通过碱基互补、配对结合,重新形成双链。在这一过程中,核酸分子经历了变性和复性的变化,以及复性过程中各分子间氢键的形成和断裂等。杂交的双方是待测核酸序列和已知核酸序列。

DNA 和 DNA 链、RNA 和 RNA 链或 2 条 RNA 链之间,只要具有一定的互补序列均可在适当的条件下发生杂交。杂交的一方是待测的 DNA 或 RNA,另一方是检测用的已知的 DNA 或 RNA 的片段,该片段则称为探针。

二、分子杂交的基本方法

分子杂交中,根据杂交中分离的对象不同,又分为 Southern 印迹杂交和 Northern 印迹杂交。鉴别 DNA 靶分子的杂交称为 Southern 印迹(Southern blot)杂交,其中待测核酸是 DNA 片段;鉴别 RNA 靶分子的杂交称为 Northern 印迹(Northern blot)杂交,其中待测核酸是 RNA 片段。由 Southern 印迹和 Northern 印迹又衍生出了斑点杂交、原位杂交和核酸酶保护等方法。

(一)Southern 印迹杂交

Southern 印迹杂交是指 DNA 与 DNA 的杂交方法。该法 1975 年由英国爱丁堡大学的 Southern 建立,故用其姓命名。该法的基本原理是:将电泳分离的待测 DNA 片段转移并结合到一定的固相支持物上,然后用标记的 DNA 探针,检测待测 DNA。主要用于测定 DNA 的限制性内切酶图谱,根据图谱可判断 DNA 的某一区域是否存在 DNA 片段的缺失、插入等基因异常,也常用于研究 DNA 的限制性内切酶片段长度多态性(RFLP);特异寡核苷酸

探针杂交还可检测到点突变，在 DNA 定量较准确的基础上，通过对杂交带上光密度扫描可进行粗略的基因定量。Southern 印迹杂交在遗传病研究、PCR 产物分析中也有十分重要的应用。

Southern 印迹杂交的基本过程见图 13－1。

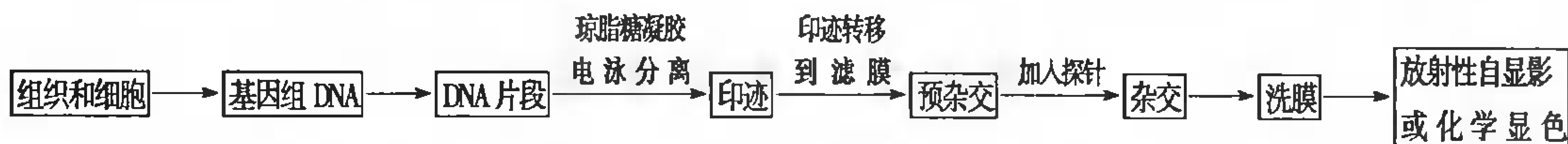


图 13－1　Southern 印迹杂交的基本过程

（二）Northern 印迹杂交

Northern 印迹杂交是指待测样品经电泳分离后转移到固相支持物上，然后与标记的核酸探针进行固－液相杂交，用以检测 RNA（主要是 mRNA）的方法。主要用于 RNA 的定性、定量研究。根据杂交图谱，可检测某种特异 RNA 是否存在、长度有无明显变化等。Northern 印迹杂交与 Southern 印迹杂交有所不同（图 13－2）。

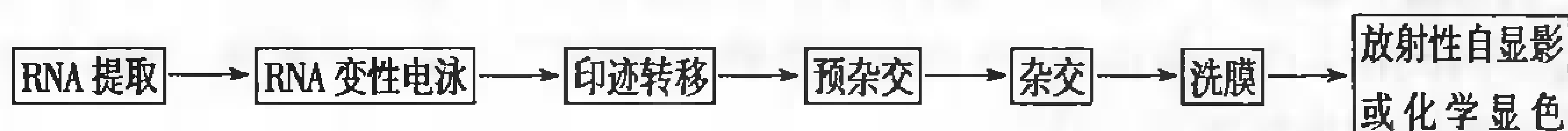


图 13－2　Northern 印迹杂交的基本过程

（三）斑点及狭缝印迹杂交

将 DNA 或 RNA 变性后直接点样于硝酸纤维膜上，用于基因组中特定基因及其表达的定性和定量研究，称为斑点印迹（dot blot）和狭缝印迹（slot blot），这也是实验室常用技术之一。与 Southern 和 Northern 印迹法相比，此技术的优点是简单、快速，可在同一张膜上进行多个样品的检测，对于核酸粗样品的检测效果也较好。缺点是不能鉴定所检测基因的分子质量，而且特异性不高，有一定比例的假阳性。

（四）染色体原位抑制杂交

在人类基因组中，大约有 20%～30%重复序列，构成相关的重复序列家族（如 *Alu* 家族，*Kpn* 家族等）。少数重复序列以串联的方式成簇分布，大部分序列散布在整个基因组内。当采用染色体文库探针或黏粒、YAC 探针进行染色体原位杂交时，那些散布于基因组中的重复序列常常会穿插出现在探针序列中，探针会与靶序列以外的同样家族成员结合，干扰了探针识别靶序列的特异性。染色体原位抑制杂交就是为解决这一难题而设计的。其基本原理是：由于 Alu、Kpn 等重复序列具有种间差异，因而，用人基因组总 DNA 作为竞争 DNA，阻断探针上的非特异性序列。用超声波将人基因组总 DNA 打成 500 bp 左右的片段，以一定比例与探针混合、变性；然后在 37 ℃条件下，孵育一段时间，这时人基因组总 DNA 和探针中的重复序列之间首先退火，此后再与染色体杂交，使探针与靶序列之间能够特异性结合。

（五）消减杂交

消减杂交（substractive hybridization）是将不同来源的基因组 DNA 或 mRNA 进行杂交，两者之间相同的核酸片段互补，或通过随机引物进行差别显示 PCR，扩增不同来源的

cDNA,比较其中的 mRNA 表达差异,从而将两种之间差异的 mRNA 筛选出来。其基本原理是:相同来源的细胞或组织中的 RNA 和 cDNA 之间可以杂交,当一种细胞中某种基因缺失或不表达,而另一种细胞没有发生变化时,把其中一种类型细胞中的单链 cDNA 与另一类型细胞中的 mRNA 或单链 cDNA 进行杂交,在两种细胞中都表达的 mRNA 就会形成 cDNA - mRNA 或 cDNA - cDNA 杂交体,而目的基因或缺失的基因则仍保持单链状态。单链 cDNA 可通过羟基磷灰石柱等方法得到分离,然后用随机引物法合成第二链或筛选 cDNA 文库,从而构建含目的基因的消减文库。

根据所选择的杂交体不同,消减杂交可分为 DNA - DNA、cDNA - mRNA、cDNA - cDNA 杂交。迄今为止消减杂交方法仍然是分离目的基因的一种有效方法。它可用于:①缺失基因片段的克隆;②不表达 mRNA 的筛选;③异常表达 mRNA 的筛选;④外源插入片段的筛选;⑤特异组织特异表达 mRNA 的筛选。

(六)比较基因组杂交

比较基因组杂交(comparative genomic hybridization, CGH)是将消减杂交和荧光原位杂交相结合的一种方法,可用于检测相关基因较小范围的缺失和扩增,并且将这些改变定位于染色体的特定区域。CGH 主要应用于肿瘤细胞遗传学的研究,其基本原理是:将肿瘤基因组 DNA 和正常基因组 DNA 分别用不同颜色的荧光染料标记,与正常中期染色体进行原位杂交,然后进行荧光分带显色,再通过检查中期染色体各个座位上两种荧光之间的荧光强度比,从而检测肿瘤 DNA 中不同序列的拷贝数的变化。

三、探针及其标记

核酸分子探针的标记,是进行核酸杂交的基础。所谓核酸分子探针是指特定的已知核酸片段,它能同与之互补的待测核酸序列退火杂交,因此可以用于待测核酸样品中特定基因序列的探测。要实现对核酸分子探针有效的探测作用,必须将探针分子用一定的标记物进行标记。根据检测对象及目的不同可选择不同的探针种类及标记方法。

(一)探针的种类

根据核酸分子探针的来源及其性质可分为基因组 DNA 探针、cDNA 探针、RNA 探针及寡核苷酸探针 4 类。

1.基因组 DNA 探针 克隆化的各种基因片段是最常用的核酸探针。由于真核基因组中存在高度重复序列,因此,探针应尽可能使用基因中外显子的序列,避免使用内含子及其他非编码序列,使用含高度重复序列的探针,可能引起非特异杂交,出现假阳性结果。

2.cDNA 探针 与 mRNA 互补的 DNA 链称为 cDNA,cDNA 中不存在内含子及其他重复序列,特异性高,是一种较理想的核酸探针。

3.RNA 探针 RNA 探针具有特异性高、本底低的特点,但由于 RNA 极易被环境中大量存在的 RNA 酶分解破坏,因此,不易操作。

4.寡核苷酸探针 根据已知的核酸序列,采用 DNA 合成仪合成一定长度的寡核苷酸片段,作为分子杂交的探针。寡核苷酸片段长度在十几个核苷酸以上就能作为杂交探针。这种探针常用于克隆筛选和点突变分析。

(二)标记物

常用的探针标记物有两类:放射性同位素标记物和非放射性标记物。放射性同位素是应用最多的探针标记物,检测具有高度的特异性,常用的有^{32}P、^{3}H和^{35}S等。放射性同位素标记物的缺点是易造成放射性污染,半衰期较短,不能长期存放。与放射性同位素相比,非放射性标记物具有安全可靠、可以长时间存放的优点。常用的有半抗原类(如:生物素、地高辛等)、荧光素(如罗丹明)等。

第三节　聚合酶链反应的原理及应用

聚合酶链反应(ploymerase chain reaction, PCR)是一种通过体外酶促反应体系,扩增特异DNA片段的方法,1985年由美国PE cetus公司的Mullis K建立。传统的DNA扩增法是分子克隆法,需经过DNA酶切、连接、转化等步骤构建含有目的基因的载体,然后导入细胞中进行扩增,再用同位素标记的探针进行筛选。操作复杂,且消耗大量时间。PCR技术的发明大大简化了分子克隆法,同时克服了传统方法的诸多不足,短时间内就可扩增获得大量的目的基因,从而比较容易对目的基因进行分析鉴定。PCR技术具有灵敏度高、特异性强、操作简便等特点,已被广泛应用于生命科学多个领域的研究中。在分子克隆、疾病诊断、遗传病分子机制的研究、法医学以及考古学等方面发挥着重要作用。PCR是20世纪末分子生物学研究领域最重大的发明之一,Mullis也因此而获得1993年度诺贝尔化学奖。

一、PCR基本原理

细胞中DNA的复制是一个复杂的过程,复制过程受多种因素影响。PCR是在试管中进行的DNA复制反应,其基本原理与细胞内的DNA复制相似,但反应体系相对简单:包括耐热的Taq DNA聚合酶、化学合成的寡聚核苷酸引物(primer)、4种脱氧核苷三磷酸(dNTP、dATP、dTTP、dCTP、dGTP)及合适的缓冲液体系。

PCR是重复进行的DNA复制过程,因而需要重复进行DNA模板解链、引物与模板DNA结合、DNA聚合酶催化新生DNA的合成(引物的延伸),这些过程都是通过控制温度来实现的,即通过改变温度引起DNA变性(denature)、退火(annealing)、延伸(extension)过程,使DNA得到复制。

1. **变性**　模板DNA在95 ℃左右高温条件下,双螺旋的氢键断裂,双链DNA解链成为单链DNA,作为反应的模板。

2. **退火**　将温度降至50 ℃左右,引物与DNA模板互补区域结合,形成杂交链。由于模板分子结构比引物复杂得多,而且反应体系中引物量较大,超过模板DNA的量,因此模板DNA单链之间互补结合机会较少。

3. **延伸**　将反应体系温度升至70 ℃左右,Taq DNA聚合酶催化以引物为起始点的5'→3'DNA链延伸反应,形成DNA链。

此3个环节为1个循环,每一个循环的产物作为下一个循环的模板,如此循环30次,引物之间的新生DNA片段理论上可以达到2^{30}拷贝,约10^{9}个分子。其原理如图13－3所示。

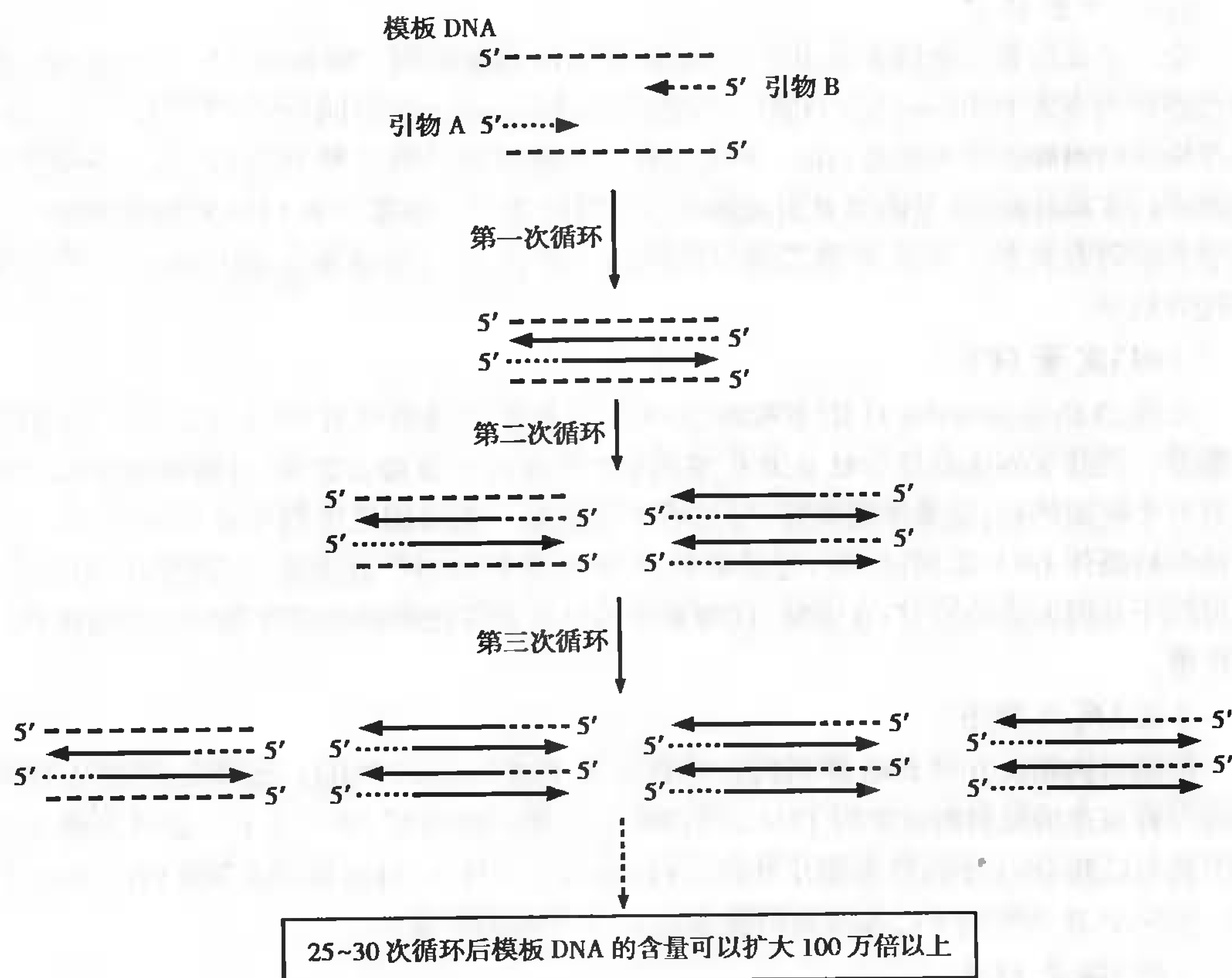


图 13－3 PCR 原理示意图

二、常用 PCR 技术

PCR 技术建立以来，因其较高的实用性而广泛应用于生命科学的各个研究领域。PCR 方法本身又在使用中不断得到发展，形成了一系列适用于不同研究目的的特殊方法。常用的有以下几种。

（一）巢式 PCR

巢式 PCR（nested－primer PCR）是指用 2 对引物先后扩增同一样品的方法。与常规 PCR 相比，其灵敏度大大提高。常用于临床检验，假阴性极少，特异性也高，临床血清、尿等样品仅需简单处理即可得到重复性很好的结果。

（二）原位 PCR

原位 PCR（*in situ* PCR）是指对组织细胞中特异 DNA 或 RNA 进行分析的一种方法。这种方法与常规 PCR 相比，不需要抽提组织细胞中的核酸，形态结构不被破坏；与原位杂交相比，灵敏度可提高 100 倍左右。目前原位 PCR 或原位杂交 PCR 主要应用于病毒的检测（主要是 HIV），肿瘤学、胚胎学以及 RNA 转运等方面的研究。

(三)多重 PCR

在一个反应管中使用多对引物,针对多个 DNA 模板或同一模板的不同区域进行扩增的过程即为多重 PCR(multiple PCR)。该技术能满足同时分析不同 DNA 序列的需要,尤其是在临床检测和科学研究时,在一个反应管中能够同时检测多种目的 DNA,这不仅能节省珍贵的实验样品,而且能使繁琐的操作变得省时省力。多重 PCR 已被成功地应用到了医学科学研究的多个方面,如各类病原的检测与鉴别,遗传病诊断及基因缺失、突变和多态性分析等。

(四)定量 PCR

定量 PCR(quantitative PCR)是指通过 PCR,定量扩增体系中的 DNA 或 RNA 的起始拷贝数量。该技术的优点是可对 PCR 扩增的全过程进行实时动态监测,可精确测定几个到几百万个起始拷贝,定量重现性好;可应用于分析某一基因的拷贝数及转录水平,也可以对特异病原体 DNA 或 RNA 进行定量研究,判断患者感染的严重程度及药物治疗的疗效;还可用于基因表达的研究;在癌症、代谢紊乱及自身免疫性疾病的诊断和分析中也具有一定价值。

(五)反向 PCR

扩增引物相反方向 DNA 序列的技术称反向 PCR(inverse PCR)。选择已知 DNA 序列内部没有切点的限制酶对此段 DNA 进行酶切,连接形成环状 DNA 分子。此时选择合适方向的与已知 DNA 序列两末端互补的引物,经 PCR 扩增后,可以得到未知序列的 DNA 片段。反向 PCR 主要用于已知序列两翼未知 DNA 序列的扩增。

(六)锚定 PCR

锚定 PCR(anchored PCR)是指用酶法在一通用引物反转录 cDNA 3′-末端加上一段已知序列,然后以此序列为引物结合位点对该 cDNA 进行扩增的方法。它可用于扩增未知或全知序列,如未知 cDNA 的制备及低丰度 cDNA 文库的构建。

(七)不对称 PCR

2 种引物浓度比例相差较大的 PCR 技术称不对称 PCR(asymmetric PCR)。在扩增循环中引入不同的引物浓度,常用(50~100:1)的比例。在最初的 10~15 个循环中主要产物还是双链 DNA,但当低浓度引物被消耗尽后,高浓度引物介导的 PCR 反应就会产生大量单链 DNA。该技术可制备单链 DNA 片段,用于序列分析或制备核酸杂交的探针。

(八)反转录 PCR

当扩增的模板为 RNA 时,先通过反转录酶将其反转录为 cDNA,然后再进行扩增的方法称反转录 PCR(reverse transcription PCR, RT-PCR)。RT-PCR 应用非常广泛,无论是分子生物学,还是临床检验等都经常采用 RT-PCR 方法。

(九)单链构象多态性 PCR

单链构象多态性 PCR(single-strand conformational polymorphism PCR,SSCP-PCR)是根据不同构象的等长 DNA 单链在中性聚丙烯酰胺凝胶中的电泳迁移率变化,来检测基因变异的技术。在不含变性剂的中性聚丙烯酰胺凝胶中,单链 DNA 迁移率除与 DNA 长度有关外,更主要取决于 DNA 单链所形成的空间构象。相同长度的单链 DNA 因其序列不同

甚至单个碱基差异所形成的构象就会不同,PCR产物经变性后进行单链DNA凝胶电泳时,每条单链处于一定的位置,靶DNA中若发生碱基缺失、插入或单个碱基置换时,就会出现泳动变位,从而提示该片段有基因变异存在。

三、PCR技术在医学领域中的应用

(一)遗传病的诊断

PCR技术已广泛应用于单基因遗传病的诊断。单基因病主要是由单个或少数核苷酸缺失、插入或置换而造成基因的不表达或表达水平低下,导致RNA加工、成熟和翻译异常或形成无功能mRNA所致。到目前为止,这些疾病还缺乏有效的治疗手段,因而,结合产前诊断进行遗传咨询、生育指导就显得尤为重要。PCR技术在遗传病产前诊断中已显示出了越来越明显的优越性。如珠蛋白生成障碍性贫血(地中海贫血)的诊断,以往珠蛋白生成障碍性贫血的产前诊断多采用胎儿脐带血、绒毛为样品,通过DNA分析进行诊断;或采集胎儿血在体外培养后,用丙烯酰胺凝胶电泳测定各珠蛋白链的合成速率比进行判断。由于这些方法需要标本量较大,成本高且耗时较长,同时部分样品只能在妊娠中期取样,具有一定的风险,因而这些方法的应用有很大局限性。而PCR技术具有成本低、快速和对样品质量及数量要求不高的特点,同时可在妊娠早期取少量样品进行操作,降低了风险性,有力地推动了产前诊断的发展。

(二)感染性疾病的诊断

利用物种间遗传基因存在特征的差别,对其进行检测是鉴定感染性疾病的病原微生物最可靠、最直接的方法,优于细菌学、生物化学和免疫学等传统的方法。已有近百种微生物可通过PCR技术检测,且大多数的检测试剂已经商品化。国内常用的有各型肝炎病毒、EB病毒、巨细胞病毒、结核杆菌、淋球菌、幽门螺旋杆菌、支原体、弓形体等检测试剂盒。PCR技术用于乙型肝炎病毒DNA(HBV-DNA)的检测,可大大提高检测敏感性,对HBV感染后复制状态及传染性的判断的认识可弥补“乙肝免疫检测五项”的不足。

PCR技术在艾滋病(AIDS)的诊断中也发挥了十分重要的作用。艾滋病是由人免疫缺陷病毒(human immunodeficiency virus, HIV)感染所致的获得性免疫缺陷综合征。对大多数艾滋病患者,可以通过血清学方法检测其血清中HIV-1抗体进行筛查并做出诊断。但在HIV-1感染早期,机体内尚无HIV-1抗体产生,以及要了解婴儿是否经母婴垂直传播途径感染了HIV-1等情况下,用PCR技术直接检测病毒则是十分必要的。

(三)肿瘤的诊断

研究发现,正常细胞基因中存在原癌基因和抗癌基因,当致癌因素作用于正常细胞,这些癌基因发生突变,使原癌基因激活或使抗癌基因失活,最终使癌基因蛋白功能异常,使细胞恶变,形成肿瘤,可以说细胞恶变的最初环节即是癌基因异常。用PCR的方法扩增癌基因片段,然后结合其他相关技术,并进行有关分析、判断是否存在碱基替换、缺失、插入等异常,从而对肿瘤进行早期诊断及治疗监测。

第四节　基因定位

基因定位(gene mapping)就是利用不同的方法将某种基因定位于染色体一定的位点上,并根据不同基因在染色体上的具体位置绘制出基因图。

人类基因定位已有近百年历史。早在 1911 年 Willson 就利用连锁分析方法第一次将人类基因——红绿色盲基因定位于 X 染色体上。但由于传统基因定位方法的局限性,基因定位的进展十分缓慢,在 1911～1967 年 56 年间定位于 X 染色体上的基因仅不到 50 个,而常染色体上的基因定位则更加艰难。1968 年,Donalae 利用系谱分析方法将 Duffy 血型基因定位于 1 号染色体上,这是人类首次将基因定位在常染色体上。20 世纪 70 年代,随着重组 DNA 技术及其他分子生物学技术的应用和发展,特别是重组 DNA 技术和细胞遗传学的结合,基因定位的方法越来越先进、准确、快捷,使基因定位研究得以迅速发展。随着人类基因组计划的提出和开展,基因定位工作已取得了丰硕成果。目前已将 4 千多个基因和 6 万余个 DNA 片段定位在特定的染色体区域。这对于研究基因与基因之间的相互关系、基因的结构与功能以及对遗传病的诊断、治疗和分子机制的研究具有重要意义。

基因定位最早是在实验的动植物中进行的,一般采用经典的连锁分析方法,鉴定有关动植物的连锁群,进一步采用细胞遗传学方法,将这些连锁群定位在特定的染色体上;还有采用杂交的方法,通过计算基因之间的重组率来确定基因之间的距离,重组率越高,两个基因之间的距离就越大。

应用连锁分析进行人类基因定位的基础是家系调查。但是,由于人类的世代时间长,家系较小,早期的连锁分析还缺乏必要的遗传标志,所以,传统的连锁方法有很大的局限性。20 世纪 90 年代,人类基因组计划的启动及一系列新方法的应用,使基因定位与克隆的研究有了飞速发展,基因定位技术也日趋完善。本节简要介绍常用基因定位的方法及其原理。

一、原位杂交

原位杂交(*in situ* hybridization,ISH)是采用固相分子杂交的方法对基因直接进行定位的一种方法。它是根据互补核酸之间能相互杂交的特性,利用标记的 DNA 或 RNA 为探针,在原位检测组织细胞内特定核酸序列的方法。这种方法的特点在于:杂交直接在载玻片的中期染色体上进行,而不是在溶液和膜上进行。所谓原位即指在标本上 DNA 原位变性,再用放射性物质(通常用^3H)标记的已知核酸探针杂交,然后通过放射自显影来检测染色体上特异 DNA 序列,用放射性颗粒在某条染色体的区带出现的最高频率或荧光强度来确定探针的位置,从而准确地进行基因定位。目前,原位杂交技术已经得到了广泛应用。例如,利用放射性标记探针对中期染色体进行杂交,已将胰岛素基因定位于 11p15。此法不足之处是,必须在已知探针的情况下方可应用此法进行基因定位,不能用此法对未知致病基因进行基因定位。

另外,也可用原位杂交技术在电子显微镜下进行基因定位。电子显微镜具有分辨率高的特性,在电子显微镜下可直接观察到核酸杂交分子。在电子显微镜下的原位杂交方

法称为 R－环法，其基本原理是，双链 DNA 分子部分变性的情况下，单链 RNA 分子和相应的 DNA 序列中的一条互补单链杂交结合并向外突出可形成环状图像；根据每种 mRNA 参与构成的 R－环的位置就可以对产生这种 mRNA 的基因进行定位，这一方法还可以用来研究断裂基因的结构。

二、体细胞杂交

体细胞杂交（somatic cell hybridization）又称细胞融合（cell fusion），是将 2 个同种或不同种细胞融合成 1 个新的杂种细胞。大多数体细胞杂交是在人、鼠（大鼠、小鼠或仓鼠）细胞间进行，在杂种细胞中，人的染色体会优先丢失。1971 年，Miller 等将小鼠 B82 细胞（胸苷激酶缺陷，TK^-）和人正常二倍体细胞（TK^+）融合，培养于 HAT 培养基中。此种培养基为一种选择系统，人和鼠细胞均不能生存在此培养基中，只有人鼠杂种细胞能存活。根据杂种细胞的生化或免疫特性检测证明，杂种细胞的存活需要胸苷激酶（TK）。经鉴定，存活的杂种细胞保留了全套的小鼠基因组，而保留有人染色体的情况则各不相同，但均含有 17 号染色体，因而可以推断人的 TK 基因定位于 17 号染色体上。

对于只存在 1 条人染色体的杂种细胞，很容易根据基因产物存在与否将相应的基因定位；对于包含几条人染色体的杂种细胞，可根据不同杂种细胞保留或丢失的人染色体重叠情况设计杂种细胞克隆嵌板（panel of clonal hybrids）进行基因定位。如表 13－2 所示，3 个杂种细胞克隆都保留有 4 条人染色体，如果某一表型特征可见于克隆 A、C，而不见于克隆 B，则表明该表型特征的基因在 3 号染色体上，因为克隆 A、C 都有 3 号染色体，而克隆 B 则无。依次类推。半乳糖激酶（GK）和尿苷－磷酸激酶（UMPK）的基因就是利用这种方法，分别定位于 17 号和 1 号染色体上的。

表 13－2　克隆嵌板

杂种克隆	保留的人染色体号数							
	1	2	3	4	5	6	7	8
A	+	+	+	+	－	－	－	－
B	+	+	－	－	+	+	－	－
C	+	－	+	－	+	－	+	－

前述的利用杂种细胞进行基因定位只是将基因定位于某一条染色体上，在此基础上，还需进行区域定位。区域定位的方法有基因剂量效应定位法、染色体易位定位法、染色体诱变定位法及统计分析定位法。

基因计量效应定位法主要是利用具有某一特定染色体的片段缺失和重复的患者或培养细胞内所观察到的基因剂量效应进行定位，如果基因产物占某一片段的拷贝数呈明显的比例变化，即可将基因定位在染色体的这一区域。

染色体易位定位法是利用杂种细胞（如人鼠杂种细胞）染色体易位后，易位染色体的对应关系进行判定。

染色体诱变断裂定位法利用人工方法在体外造成杂种细胞的特定染色体不同区域发生断裂，形成各种类型的缺失，得到一套特定染色体缺失的杂种细胞，再根据基因产物进

行定位。

统计分析定位法的基础是靶学说:2个基因之间距离越大,这2个基因由于染色体断裂而分开的可能性越大。利用X射线、γ射线等方法处理人鼠杂种细胞,测定一系列基因产物和同工酶,统计分析构成基因分离的数据,从而判断基因之间的距离。

三、荧光原位杂交

荧光原位杂交(fluorescence *in situ* hybridization,FISH)是以荧光标记取代同位素标记而形成的一种原位杂交方法。探针首先与某种标记物分子(如生物素或地高辛等)结合,杂交后通过免疫细胞化学过程连接上荧光染料,在荧光显微镜下观察。这种方法具有高度特异性,有与放射性探针同样甚至更高的分辨率。目前,中期染色体FISH的最大分辨率为几个Mb,若用前中期染色体,分辨率可达1 kb。人染色体的DNA线性长度平均为5 cm,经多级螺旋化和折叠形成的中期染色体长度大为缩短,因此,原位杂交对高度螺旋化的中期和前中期染色体不能得到高分辨的制图。要获得高分辨率可通过2种方式:一是通过延长中期染色体,人为拉长DNA纤维,即在载玻片的一端用清洁剂溶解细胞,倾斜载玻片,使溶剂中的DNA沿玻片流下,可使分辨率达到1至几个kb;二是通过培养细胞的同步化,使大量细胞处于有丝分裂前期来获得高分辨率。如将间期核用于杂交,可进一步提高分辨率,间期FISH技术能对间隔只有50 kb的探针排序。新近应用的精－卵系统能使FISH技术的分辨率达到小于50 kb的范围。

近年来,用不同荧光染料进行多重原位杂交,标本显示不同的荧光色泽,这种多色荧光技术发展迅速,已由常用的双色、三色已发展为多色荧光原位杂交。目前,荧光技术与计算机软件相结合,已可同时鉴定27种不同的DNA杂交探针,能迅速准确地发现多种临床标本和肿瘤细胞系的细胞内染色体的简单或复杂的染色体重排、缺失,显示了多重荧光原位杂交技术不仅可以用于基因定位,而且对复杂核型的分析有很大优势,使细胞遗传学具有更大的临床应用价值。

四、连锁分析

连锁分析(linkage analysis)是通过基因与遗传标记之间的重组率来估计基因在染色体上的位置,是最早采用的基因定位方法。用于连锁分析的遗传标记(另一个已知基因)是基因定位的基础,遗传标记越多,杂合性越强,基因定位就越精确、越容易。遗传标记的选择经历了经典遗传标记(如红细胞ABO血型、白细胞HLA座位标记等)、第1代遗传标记(限制性片段长度多态性,RFLP)、第2代遗传标记(短串联重复序列,STR)、第3代遗传标记(单核苷酸多态性,SNP)的发展过程。连锁分析的方法也随之得到相应改进。

(一)RFLP家系连锁分析

1980年,Botsftin等将限制性片段长度多态性(restriction fragment length polymorphism,RFLP)方法用于基因诊断,标志着基因定位研究进入了一个新的时期。RFLP是指利用限制酶将人群中等位基因DNA片段多态变化呈现出来。RFLP广泛存在于人类基因组中,呈孟德尔式遗传,而且许多已经定位,因而可以用于连锁分析和基因定位。应用RFLP进行基因定位的优点在于:通过RFLP与疾病基因的连锁关系,可直接分析RFLP与疾病基

因的连锁性，不需要分析基因产物，拓宽了基因定位的范围。目前通过 RFLP 方法定位的基因已达到 400 余个，如亨廷顿舞蹈病基因、囊性纤维化基因及成人多囊肾基因等。

（二）STR 连锁分析

1985 年，英国 Leicesfer 大学的 Jeffreys 及其同事在人肌球蛋白基因中发现了一些短的简单重复单位，即“小卫星中心（minisatellite core）”。随后，许多研究证明，在人类基因组中分布有大量的此类长度的多态性标记，它们串联成簇，分布于基因组的各个位置。在某一点上，数量可变的串联重复（variable number of tandem repeat，VNTR）可提供不同长度的片段，其重复单位长度为 6 ~ 12 bp。1989 年，另一类称之为“微卫星标记（microsatelite marker）”系统被发现和建立，其重复单位长度为 2 ~ 6 bp，因而又被称为短串联重复（STR）。STR 具有高度多态杂合性，也符合孟德尔遗传方式，在人类基因组中广泛存在，约占真核生物基因组的 5%。目前发现的 STR 位点超过 8 000 个，平均 100 ~ 200 kb 就有 1 个 STR 位点。同时，PCR 与 STR 位点分析的结合，操作简便、容易自动化，而且不受探针的限制。因而，STR 作为遗传标记已取代 RFLP 应用于基因定位研究中。目前，应用 STR 遗传标记定位并克隆的基因有 Wilms 瘤基因、同性恋相关基因等，推动了多基因病易感基因定位的研究。例如，采用 STR 遗传标记，应用全基因组扫描技术，发现了 12 个基因位点与 I 型糖尿病高度相关，其中包括 *HLA* – Ⅱ类基因位点。

（三）SNP 连锁分析

随着人类基因组研究的发展，1996 年美国麻省理工学院的人类基因研究中心负责人 Lander 又提出了一类称为 SNP 的新的遗传标记系统，即单核苷酸多态性标记。SNP 在人群中有相当大的遗传多态性，而在大多数基因位点上都会有若干个等位基因；对每一个核苷酸来说，其突变率大约为 10^{-9}，这就意味着每一个核苷酸在任何一代人群中大约每 1×10^{9} 个个体就会发生一次变异。由这种方式产生的点突变就形成许多双等位型标记，这种标记在人类基因组中可达 300 万个，平均每 1 000 个碱基对就会有 1 个。因此，在基因定位研究中，SNP 遗传标记系统有着其他标记系统不可比拟的优越性和潜力，成为目前基因定位的主要遗传标记。

（四）基因定位相关的连锁分析方法

在用于基因定位的 DNA 标记研究取得进展的同时，基因定位的相关连锁分析方法也有了长足的发展。根据基因的重组率来计算两基因之间的染色体图距称为连锁分析。近年来，用于连锁分析的计算机软件的开发，使复杂的计算分析方法应用得越来越广泛。

1. 参数型连锁分析方法——优势对数计分法 1955 年，Morton 利用似然性估计的原理提出了优势对数分数法（log odds score），简称为 LOD 法。LOD 法主要检测在两基因以某一重组率（θ）相连锁时，出现这种情况的似然性（L）有多大。LOD 法的基本公式为：

$$Z(\theta) = \log 10 \frac{L(\theta)}{L(0.5)}$$

$L(\theta)$表示重组率为 θ 时连锁的可能性。$L(0.5)$表示重组率为 50%，即不连锁的可能性。这 2 个可能性之比的对数就是连锁优势的对数计分。如 LOD 值 > 1 支持连锁，LOD 值 > 3 肯定连锁，LOD 值 < – 2 否定连锁。重组率的判断为：$\theta < 0.10$ 为紧密连锁（遗传距

离 < 10 cM)，$\theta > 0.20$ 为松弛连锁(遗传距离 > 20 cM)，$0.10 < \theta < 0.20$ 为中度连锁。

LOD 法的优点在于：①可用于 2 代小家系资料的分析；②每个家系的计算结果可逐个相加，无时间限制，一旦能做出判断即可停止调查；③计算基因连锁的似然性及染色体图距准确性较高；④结果判断较容易。

LOD 法的不足在于：不适合 3 代或 3 代以上大家系的连锁分析。另外，亲代基因型必须已知，且其中必有 1 个为双重杂合子。

LOD 法是基于一定的遗传假设之上的，包括基因频率、外显率、遗传方式等，故又称参数性或模式依赖的连锁分析方法。这种方法的效力依赖于所选择的遗传方式是否正确，所用标记物的信息含量、家系大小、患病人数及遗传异质性、外显率和基因频率等多种因素。可用于连锁分析的计算机软件包主要有 Liped、Linkage 和 Vitesse 等。

该法主要适用于已知遗传方式的单基因遗传病的基因定位，如外显率完全的显性遗传病和能定量的隐性遗传病。对于多基因遗传病易感基因的定位，该法的作用非常有限。其原因是：首先，多基因遗传病的遗传没有一个固定的模式，无法设置参数并套用公式；其次，多基因遗传病群体发病率相对较高，基因突变的频率增大，影响了基因定位时目的基因的确定；另外，多基因遗传病在不同家系中，甚至在一个大家系中存在遗传异质性，使这些家系资料无法累加。非参数型连锁分析方法则为解决这一难题开辟了有效的途径。

2.非参数型连锁分析方法　非参数型连锁分析方法包括受累同胞对分析、受累家系成员法和传递不平衡检验。

受累同胞对(affected sib－pair，ASP)分析法是连锁分析的一种特殊形式，其特点是无需知道遗传病的遗传方式，即可对同胞对中某一遗传标记与疾病易感基因做出连锁关系的判断。在哮喘病、精神分裂症、原发性高血压和老年性痴呆等多基因遗传病的研究中，应用 ASP 法已找到了这些疾病的一些易感基因或基因位点。ASP 法的不足之处在于：为了排除遗传异质性的影响，往往需要收集几百个受累同胞对，这对一、两个实验室来说是很难做到的；另外，不能像 LOD 法那样得出遗传标记和易感基因之间的距离。

受累家系成员 (affected pedigree member，APM) 法是一种与 ASP 相类似的方法。它的原理与 ASP 法相同，只是把分析对象扩展到整个家系的所有成员，从而解决了 ASP 法分析时家系资料不足的问题。

APM 法与 ASP 法不同之处在于：分析家系中所有个体的状态同一(identical by state，IBS)，即只考虑家系成员的遗传标记或等位基因的相似性，而不考虑其是否来自于共同的祖先。因此，该法取材容易，但分析遗传标记和易感基因之间的有效性要比 ASP 法低。目前，APM 法较多用于同胞对收集较困难的晚发性多基因遗传病的连锁分析。如该法用于晚发型 Alzheimer 病的研究，确立了 Apo E 基因与该病的连锁关系，并将 Apo E 基因定位在 19 号染色体上。

传递连锁不平衡检验(transmission disequilibrium test，TDT)法原理是：杂合子亲代的任何 DNA 多态位点的 2 个等位基因传递给子代的概率应是随机的(即各为 0.50)，如果子代患者从亲代传得的等位基因频率显著偏离 0.50，则该标记可能与疾病或其中间性状位点有关。应用 TDT 法不受群体分层的影响，检出率高于连锁分析，可同时进行性状的定性和定量检验。应用此法需要测定患者及其父母的遗传标记基因型，对晚发的多基因遗传

病而言,收集患者的双亲标本有一定困难。

五、定位克隆与定位候选克隆

(一)定位克隆

定位克隆(positional cloning)是指通过家系分析或缺失分析,将某种遗传性疾病或遗传性状的决定基因定位在某个染色体的特定区域,然后不断缩小候选区域,最后鉴定出这个基因并研究其功能的方法。自1986年以来,随着大片段DNA克隆技术的发展,定位克隆技术的应用越来越广泛。通过这种方法人们已克隆分离出了慢性肉芽肿、假性肥大性肌营养不良症和亨廷顿舞蹈病等疾病的几十个致病基因。

定位克隆的基本方法是:①通过染色体缺失、平衡易位或连锁分析(特别是DNA连锁分析),确定所研究基因在染色体上的位置,并将这个位置精确到2 000 kb左右;②确定YAC文库,利用距离这个基因最近的DNA标志筛选YAC文库,然后用染色体步移技术获得覆盖这个基因位点的一组连续的YAC克隆;③寻找基因,在确定的YAC文库区域内寻找基因;④鉴定基因,首先要对寻找到的基因所编码的蛋白质氨基酸序列进行分析。如果是一个致病基因,应分析患者群体中该基因是否存在DNA片段缺失、重排以及突变等结构变化,还应分析这些改变是否存在规律性。

(二)"定位-候选"克隆

"定位-候选"克隆(positional candidate cloning)方法是随着将致病基因定位到一个小的染色体区域而产生的方法。所有定位在这样区域内的基因不仅根据其蛋白质产物的特征还要根据其在图谱中的位置来确定。

当一个新的致病基因定位在染色体的某一区带后,可以将计算机数据库中同样定位在这一染色体区域上的基因或cDNA的功能结构域、序列稳定性、表达的时空特异性以及发育与遗传对表达的影响等资料调出,与这一疾病的生化改变、受累组织、遗传特征、发育缺陷和动物模型等进行分析比较,找出候选基因,并进一步在患者个体中筛查该基因是否发生结构异常,以及这种结构改变是否导致功能改变,从而确认该候选基因是否是致病基因。

小　结

人类基因研究的主要技术包括:重组DNA技术、PCR技术、分子杂交技术和基因定位技术等。

重组DNA技术是基因工程的核心技术,对遗传病的诊断以及人们认识遗传病、免疫性疾病和肿瘤的分子机制具有重要意义。它的基本过程包括:①制备目的基因和相关载体;②将目的基因和有关载体进行连接,形成重组体;③将重组的DNA导入受体细胞;④DNA重组体的筛选和鉴定;⑤DNA重组体的扩增、表达和其他研究。在重组DNA技术中,工具酶和载体的选择是关键环节。工具酶包括各种有特异性剪切位点的限制酶及修饰酶两类;根据不同的实验目的需要选择不同的载体,常用的载体有:质粒、噬菌体、酵母人工染色体(YAC)和病毒等。

PCR 是在试管中进行的 DNA 复制反应。模板 DNA 在 PCR 反应体系中，通过变性、退火和延伸 3 个步骤反复循环，每次循环的产物作为下一循环的模板，n 次循环后，可使模板 DNA 分子扩增到 2^n 倍。在 PCR 基础上形成的适用于不同研究目的的 PCR 技术系列包括：巢式 PCR、原位 PCR、多重 PCR、定量 PCR、反向 PCR、锚定 PCR、不对称 PCR、反转录 PCR 和 SSCP - PCR 等。

分子杂交是定性和定量检测特异 DNA 或 RNA 序列片段的有力工具，其基本原理是：将标记的核酸探针与待测核酸片段混合在一起，经过核酸变性、复性后，如待测核酸中含有与探针互补的片段，会与探针配对结合，重新形成双链，从而达到检测目的。检测不同的核酸分子需用不同的杂交方法。鉴别 DNA 靶分子的杂交称为 Southern 印记杂交；鉴别 RNA 靶分子的杂交称为 Northern 印记杂交。在此基础上，随着分子遗传学的发展，又产生了许多新的杂交方法，如荧光原位杂交、比较基因组杂交和消减杂交等。

基因定位是研究遗传病发病机制、进行疾病相关基因的分离与克隆以及基因诊断的基础。采用连锁分析方法进行基因定位，这是基因定位研究中应用最广泛的方法。连锁分析首先要有合适的遗传标记，最初用于连锁分析的遗传标记是 ABO、HLA 等多态性位点。随着分子遗传学的发展，DNA 标记在基因定位中被广泛应用，从 RFLP、STR 到 SNP，多态性信息量越来越大，基因定位也越来越精细。另外，连锁分析的方法也在不断发展。

（霍正浩）

第十四章　遗传病的诊断与治疗

遗传病的诊断(diagnosis of hereditary disease)是一项复杂的工作,需要多学科的密切配合,是遗传病的预防和治疗的前提。近年来,随着重组 DNA 技术在医学领域的广泛应用,遗传病的治疗有了长足的进步,从传统的手术治疗、药物治疗、饮食疗法等发展到基因治疗。随着技术、方法的日臻成熟,遗传病的根治性治疗必将掀开新的一页。

第一节　遗传病的临床诊断

遗传病的临床诊断是指从病史入手,通过患者叙述的症状、体格检查获得的体征及系谱分析、皮纹分析的结果,对患者所患疾病是否为遗传病做出初步判断。

一、病史、症状和体征

遗传病的病史采集与非遗传病类似,但有时患者或代诉人由于文化程度,记忆、思维及判断能力的局限,可使病史、症状的描述不够准确或不全面;患者或代诉人因某些原因可能提供假病史,应予以鉴别。由于遗传病多有家族聚集现象,除一般病史外,应着重询问患者的家族史、婚姻史和生育史。

家族史中,应注意询问有无遗传病患者及发病情况;对婚姻史,要注意询问婚龄、配偶健康状况以及是否近亲结婚;对生育史,要注意询问生育年龄、生育子女数及健康状况,有无流产、死产或早产等,患儿出生时有无产伤、窒息,妊娠早期有无病毒性疾病或接触过各种致畸因素等。

症状或体征是患者就诊的主要原因,也是遗传病诊断的重要线索。遗传病既有与非遗传病相同的症状和体征,又有与特定遗传病相关的特有症状和体征。特有症状和体征是遗传病诊断的初步线索。例如,智力发育不全伴特殊腐臭尿味提示苯丙酮尿症,智力发育不全伴白内障、肝硬化等提示半乳糖血症,智力发育不全伴生长发育迟缓、五官及四肢畸形提示常染色体病,外生殖器发育异常、第二性征发育落后或异常、原发性闭经等提示性染色体病。

二、家系调查与系谱分析

对患者进行家系调查和系谱分析,有助于了解患者所患疾病是否属于遗传病,是单基因遗传病还是多基因遗传病;如果是单基因遗传病,则可分析出属于何种遗传方式。家系调查时,从先证者入手,调查先证者亲属患病情况,将调查材料绘制成系谱图,根据系谱中家系成员的婚配和发病情况,对患者所患疾病做出判断。

系谱分析时应注意以下问题:①系谱的系统性、完整性和可靠性。完整的系谱至少应包括 3 代家族成员;家族成员要逐个查询,关键成员更不可遗漏;死亡者(包括婴儿死亡)

需查清死因;死胎、流产也要纪录在系谱中。②有的显性遗传病具有外显不全和延迟显性的特点,要注意与隐性遗传病区分。③在有的家庭中,遗传病是新产生的突变基因所致的,不可误认为是隐性遗传病。有些遗传病家系中,除先证者外,家庭成员中找不到其他患者,因而很难从系谱中判断其遗传方式,不能轻易依据家系中的"散发"病例确定为常染色体隐性遗传病。例如,约 1/3 的 X 连锁隐性遗传的假性肥大性肌营养不良症患者为新的基因突变引起。④显性与隐性的相对性。同一遗传病由于判断遗传方式所依据的观察指标不同,可能判断为不同的遗传方式。例如,在临床水平,镰形细胞贫血症纯合子有严重的贫血,正常情况下杂合子无贫血,因此,以贫血作为观察指标,突变基因被认为是隐性的;在氧分压低的情况下,杂合子也有少量红细胞形成镰刀状,所以在细胞水平,以镰形红细胞数量为观察指标,突变基因呈不完全显性,杂合子(HbAHbS)镰形红细胞数量介于正常纯合子(HbAHbA)与突变基因纯合子(HbSHbS)之间。

三、皮肤纹理分析

人体的手、脚掌面具有的皮肤纹理是妊娠 14 周时形成的。掌面皮肤真皮向表面突出形成嵴,嵴与嵴之间形成沟,嵴和沟组成皮肤纹理,称皮纹。皮纹的形成是遗传和环境共同作用的结果,因此,皮纹可作为诊断某些遗传病的辅助手段。用于分析的皮纹包括指纹、掌纹、掌褶纹和足纹,其中以指纹和掌纹的研究较多。

指纹可分为弓形纹(arch,A)、箕形纹(loop,L)和斗形纹(whorl,W) 3 种类型。弓形纹的纹线从指的一侧发出到另一侧;箕形纹的纹线从一侧发出后中途转回;斗形纹的纹线呈同心圆状或螺旋状排列。其中 3 个方向走行嵴线的交汇处,称三叉点(triradius)。弓形纹没有三叉点;箕形纹有 1 个三叉点,按开口方向分为尺箕(向尺侧)和桡箕(向桡侧);斗形纹有 2 个三叉点(图 14-1)。



图 14-1　指纹的几种类型

掌纹可分为 5 个部分:①大鱼际区(thenar area),位于拇指下方;②小鱼际区(hypothenar area),位于小指下方;③指间区(interdigital area)($Ⅰ_1$ - $Ⅰ_4$);④第 2、3、4 和 5 指基部,各有 a、b、c 和 d 三叉点,并引出主线 A、B、C 和 D;⑤t 三叉点,位于手掌基部。从 a 三叉、d 三叉分别至 t 三叉连线,两线的夹角即为 atd 角(图 14-2)。

正常人手掌褶纹有 3 条:远侧横褶纹、近侧横褶纹、大鱼际褶纹。如果远侧横褶纹和

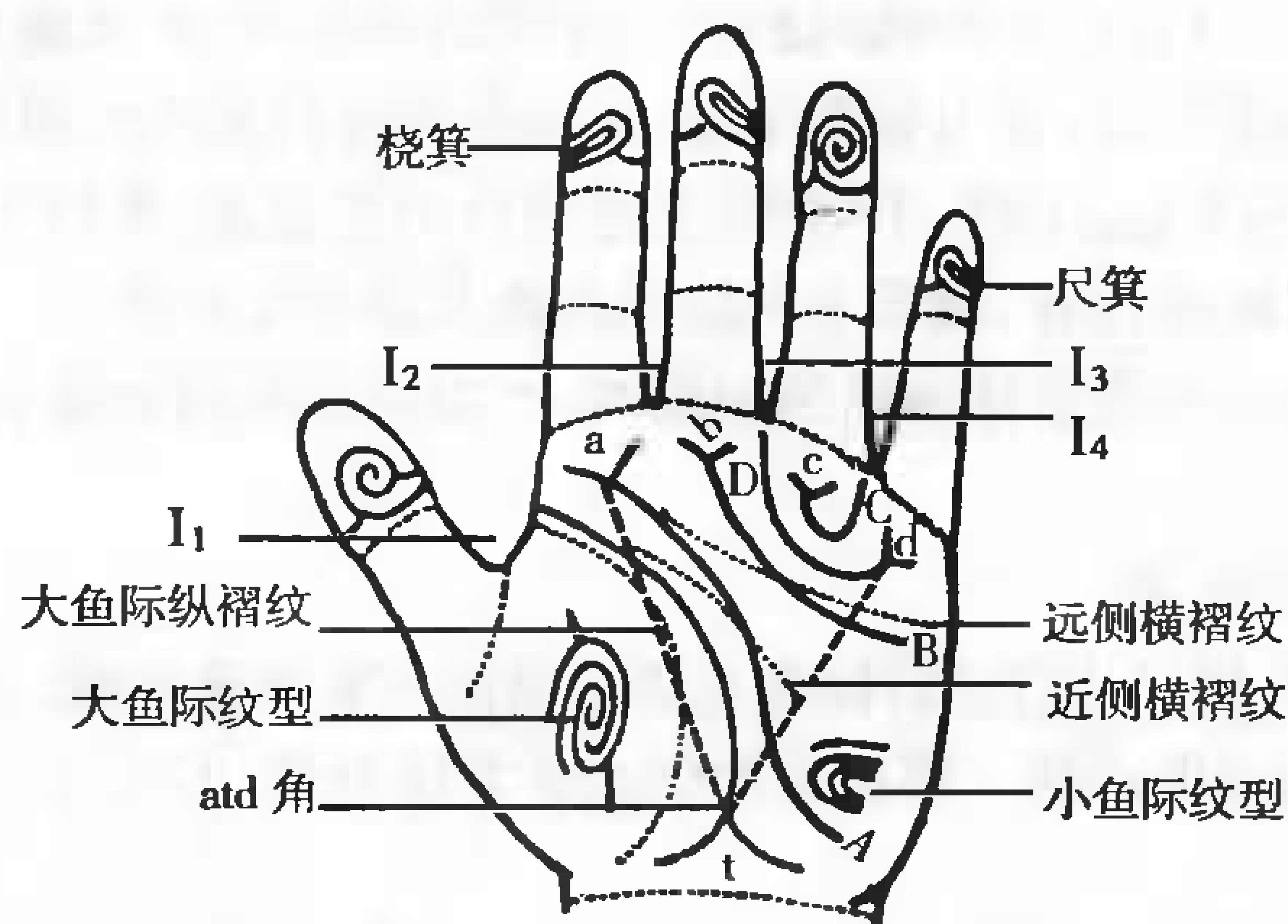


图 14－2　手掌主要掌纹线

近侧横褶纹连接成一条直线，从桡侧至尺侧呈水平横贯手掌，则称为通贯手(图 14－3)。

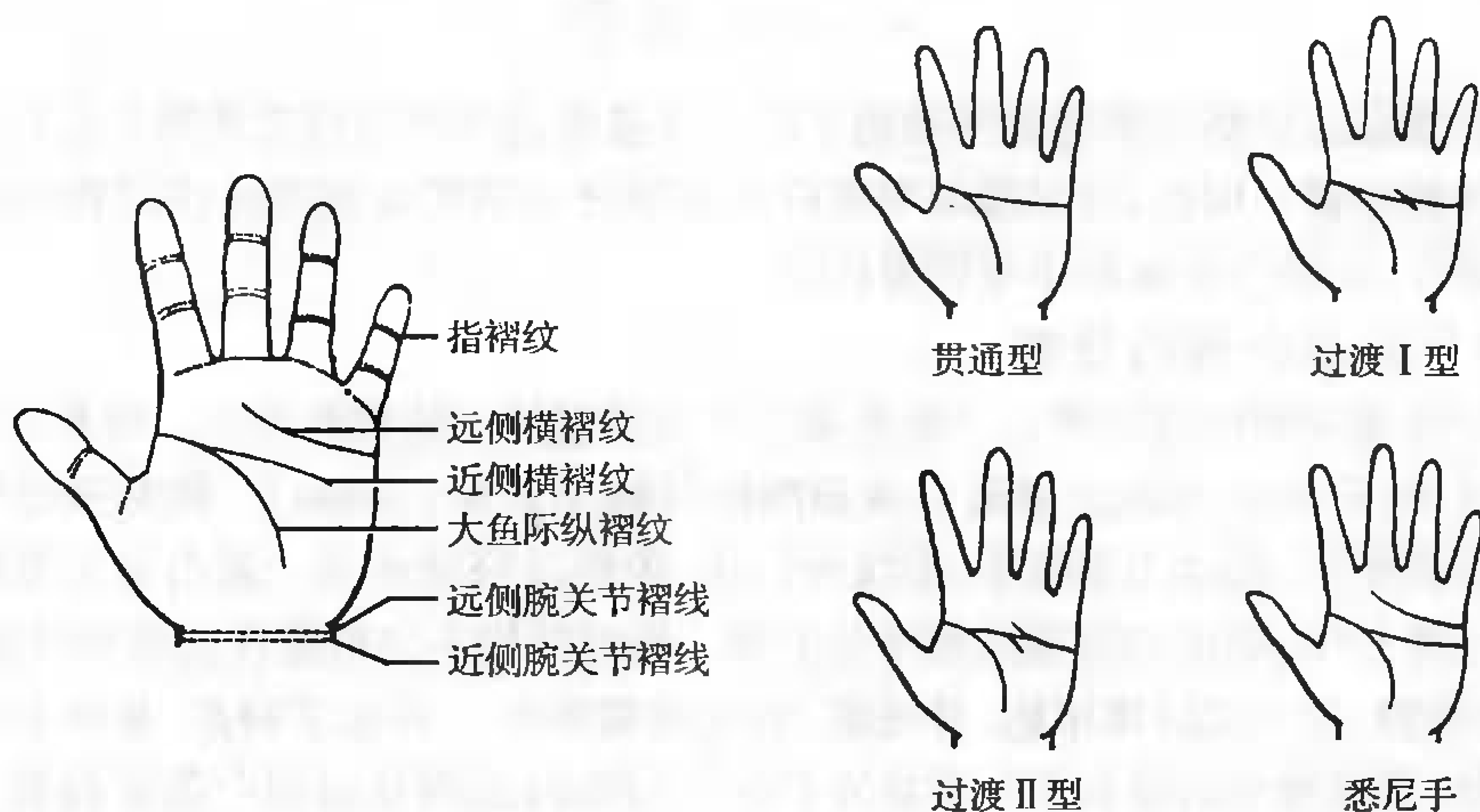


图 14－3　指褶纹、掌褶纹及其变异类型

遗传病，特别是染色体病，患者的皮纹往往出现某些特定组合。例如，21 三体综合征患者的指纹以尺箕为多；atd 角常大于 45°；约 1/3～2/3 患者有通贯手；其他染色体病，如 18 三体综合征、13 三体综合征、5p－综合征、Turner 综合征和 Klinefelter 综合征等患者的皮纹都有某些特征性改变。

第二节　遗传病的实验诊断

一、细胞遗传学检查

(一)染色体检查

染色体检查即核型分析，是确诊染色体病的主要方法。随着显带技术的发展以及高

分辨显带技术的应用,目前已能准确判断染色体的结构异常,甚至能发现染色体微畸变。

染色体检查的指征包括:智力发育不全、生长迟缓伴其他先天畸形患者,习惯性流产、早产、死产等不良妊娠夫妇,不孕、不育夫妇,原发性闭经患者,无精子症患者,性征发育不全或两性内外生殖器畸形患者,疑为先天愚型的患儿及其父母等。

染色体检查大多取外周血做淋巴细胞培养,产前诊断时则取绒毛、羊水中胎儿脱落细胞、脐带血等。

(二)性染色质检查

性染色质(X 染色质、Y 染色质)检查主要用做性发育异常疾病、真假两性畸形及性染色体数目异常疾病的辅助诊断。该技术的优点是方法简单,但不完全可靠,确诊需进行染色体检查。

性染色质检查的标本取自口腔黏膜上皮细胞、头发根鞘细胞、皮肤、女性的阴道上皮细胞,以及绒毛和羊水的胎儿脱落细胞等,标本涂片染色后进行观察;检测 Y 染色质时,标本需用盐酸喹吖因染色,然后在荧光显微镜下观察。

二、生化检查

生化检查是遗传病诊断的重要辅助手段。单基因遗传病的病变基础是蛋白质或酶分子结构或功能异常。因此,可以通过对蛋白质或酶分子结构或酶促反应过程中的底物或产物进行定性、定量分析诊断单基因遗传病。

(一)蛋白质和酶的检测

蛋白质含量和酶活性的测定是确定某些单基因遗传病的重要方法。随着生化检测技术的不断进步,目前还可鉴别出蛋白质和酶的结构变异型(variant)。酶变型的分析方法主要有电泳速率法、酶动力学技术、指纹分析法、免疫反应技术等。蛋白质变型的分析方法主要有电泳技术、肽链和氨基酸顺序分析等。检测的标本包括患者血液和特定的组织、细胞,如肝细胞、皮肤成纤维细胞、肾细胞、肠黏膜细胞等。需注意的是,基因表达具有特定的时空性,即某种酶的异常不一定在发育的任何阶段及所有组织中都能检测到,例如,检测苯丙氨酸羟化酶必须用活体肝组织,在血细胞中则检测不到。

(二)代谢产物检查

酶的异常导致一系列代谢紊乱,代谢中间产物、底物、终产物或旁路代谢产物等发生质和量的变化。因此,检测某些代谢产物的质和量的改变,可间接反映酶的变化,从而对疾病做出诊断。例如,苯丙酮尿症患者血清苯丙氨酸浓度增高,尿苯丙酮酸、苯乳酸、苯乙酸含量增加;检测血清磷酸肌酸激酶活性可对假性肥大性肌营养不良症做出诊断;尿硫酸皮肤素、硫酸乙酰肝素含量变化有助于黏多糖病的诊断。

三、基因诊断

基因诊断(gene diagnosis)就是利用 DNA 重组技术直接从 DNA 水平上检测人类遗传病的基因缺陷,它和传统的诊断方法主要差别在于直接从基因型推断表型,即可以越过基因产物直接检测基因结构而做出诊断。基因诊断具有针对性强、特异性强、灵敏度高、诊断范围广、目的基因无组织和发育特异性等特点。然而,由于基因突变的多样性,除了缺失、

插入、倒位、动态突变和一些高发的点突变外，大多数基因突变难以直接检测。因此，目前的基因诊断主要是以先证者为线索进行产前诊断。

（一）RFLP 连锁分析法

点突变可导致基因上限制性酶酶切位点的消失或新的酶切位点出现，不同个体基因组 DNA 用同一限制性酶切割时，DNA 片段长度出现差异，这种限制性酶酶切位点变化所导致的 DNA 片段长度的差异，称为限制性片段长度多态性（RFLP）。RFLP 可用 Southern 印迹杂交法检出。

根据待测个体的限制性酶酶切片段的大小和数量变化，可做出是否存在基因突变的诊断。例如，已知镰形细胞贫血症的突变基因是 β^S 基因，即 β^A 基因第 6 位密码子由 GAG 变为 GTG 所致。限制性酶 MstⅡ的识别序列是 CCTNAGG，能识别并切割正常 β^A 基因的 CCTGAGG 序列，但不能识别、切割 β^S 基因的 CCTGTGG（A→T）序列。因此，镰形细胞贫血症由于基因突变（β^S 基因）少了一个识别位点。用 β^A 基因探针对基因组 DNA 的 MstⅡ酶切产物进行 Southern 杂交，正常人（$\beta^A\beta^A$）有 1.15 kb 片段和 0.2 kb 片段，患者（$\beta^S\beta^S$）只有 1.35 kb 片段，而杂合子（$\beta^A\beta^S$）有 1.15 kb、0.2 kb 和1.35 kb 3 个片段（图 14－4）。

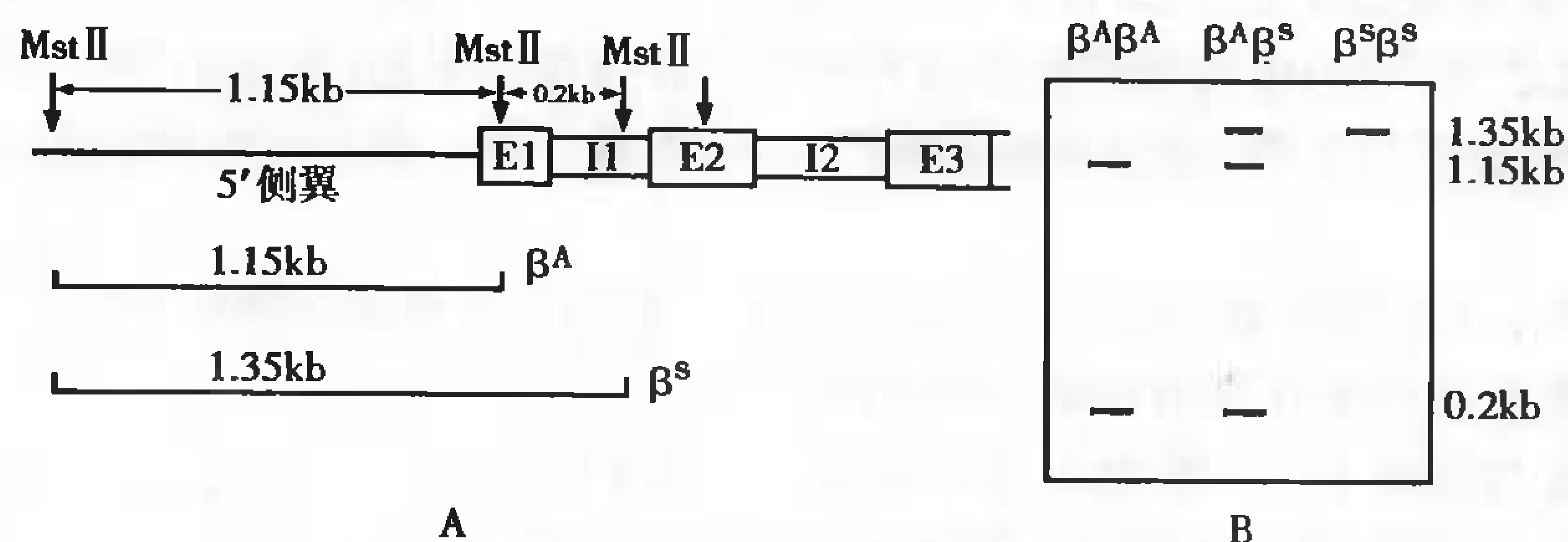


图 14－4 镰形细胞贫血症的基因诊断
A：β珠蛋白基因 5′侧翼区 MstⅡ酶切位点模式图
B：β^A 基因和 β^S 基因 DNA MstⅡ酶切产物电泳示意图

（二）STR 多态连锁分析法

短串联重复（STR）是存在于人类基因组中的重复序列，重复单位为 2～6 bp，也称为微卫星 DNA，因在人群中重复次数不同而呈多态性，它的杂合度和所含信息量比 RFLP 高。因此，可通过 PCR 法检测 STR 的多态性，再通过连锁分析进行基因诊断。

例如，肝豆状核变性为 AR 病，致病基因位于 13q14，在该基因旁侧的 $D_{13}S_{301}$ 基因位点为 STR，人群中共有 8 个等位片段（A1～A8）。图 14－5 为一肝豆状核变性家系的 $D_{13}S_{301}$ 位点连锁分析图，$Ⅰ_1$ 的 A_6、$Ⅰ_2$ 的 A_3 与致病基因 d 连锁，$Ⅱ_3$ 与 $Ⅱ_2$ 基因型相同，但没有发病，诊断为症状前患者。

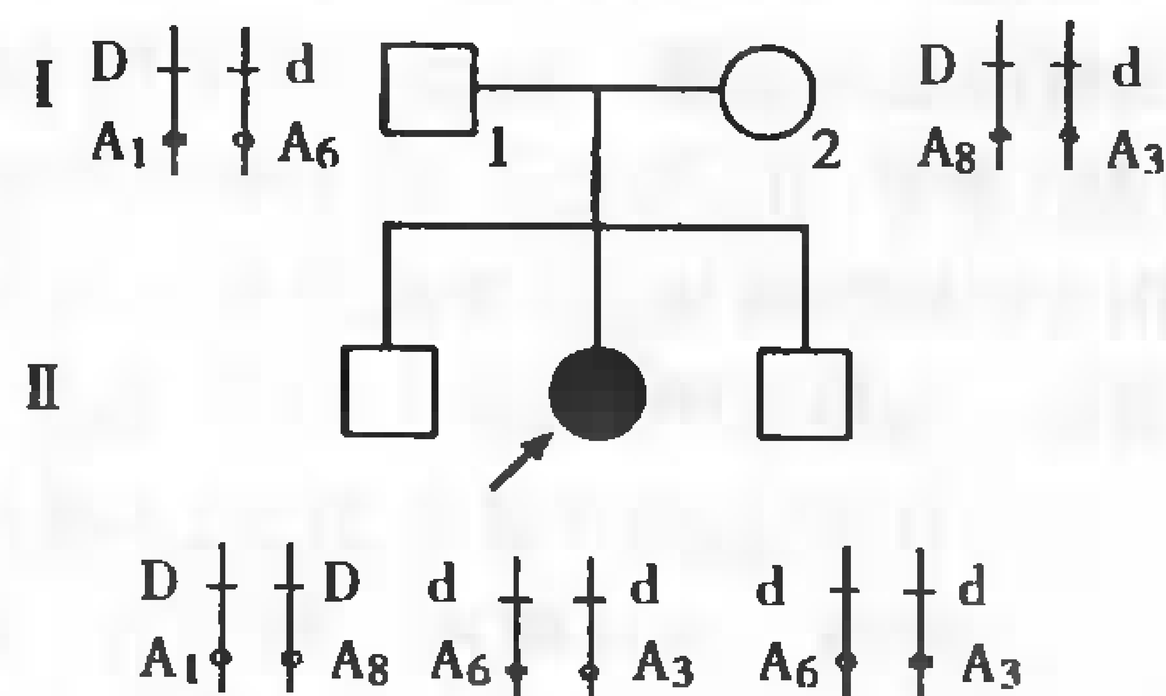


图 14－5 肝豆状核变性家系的 $D_{13}S_{301}$位点连锁分析

(三)ASO分析法

当基因突变部位和性质完全明了时,可以应用同位素或非同位素标记的等位基因特异的寡核苷酸(allele - specific oligonucleotide,ASO)探针进行诊断。探针长度通常为20 bp左右。检测点突变时一般需要合成2种探针,一种与正常基因序列完全一致,能与之稳定杂交,但不能与突变基因序列杂交;另一种与突变基因序列一致,能与突变基因序列稳定杂交,但不能与正常基因序列杂交。这样,就可以把只有1个碱基突变的基因区别开来。

例如,应用ASO分析法诊断镰形细胞贫血症时,需分别合成β^A基因探针和β^S基因探针,如果待测基因为β^S基因,则不能和β^A基因探针杂交结合,而可以和β^S基因探针杂交结合,从而对待测基因做出诊断。

ASO探针分析法还可以和PCR技术结合,即先以PCR技术扩增长约110 bp的基因片段,然后再与ASO探针杂交,这样,可以减少杂交时的非特异性信号。

(四)PCR技术

许多单基因遗传病可通过PCR技术诊断。如果疾病由基因缺失引起(如α珠蛋白生成障碍性贫血),可在缺失两端设计一对引物,扩增结果是无扩增产物,或是缩短了的扩增产物;如果疾病由点突变引起,且突变的位置和性质已知,可设计包括突变部位的引物,扩增结果也无扩增产物;也可根据突变位置和性质设计相应引物(突变引物),结果是:用正常引物扩增,无扩增产物;而用突变引物扩增,有扩增产物。据此可对突变的存在做出诊断。

PCR技术在发展过程中衍生了一系列适用于不同研究目的的特殊PCR技术,以下简介两种特殊PCR技术在基因诊断中的应用。

1.多重PCR技术　多重PCR技术可在一个反应体系中加入多对引物,同时扩增多个部位的DNA片段,引物的扩增温度相同,但扩增的DNA片段大小不同,电泳时可完全分开。例如,可应用多重PCR技术诊断缺失型DMD。

DMD的遗传方式为XR,DMD基因很大,约70%的DMD突变为发生在不同部位的缺失型突变。经多重PCR检测,如发现扩增产物的电泳条带缺失,即可对缺失做出定位诊断。在图14-6的DMD家系中,II_1为DMD患者,II_2为发育50天的男性胚胎,取其绒毛组织作为扩增的模板,其他成员取外周血白细胞作扩增的模板。结果表明,患者II_1 DMD基因有外显子50、52缺失,II_2 DMD基因外显子50、52无缺失,为正常男性。

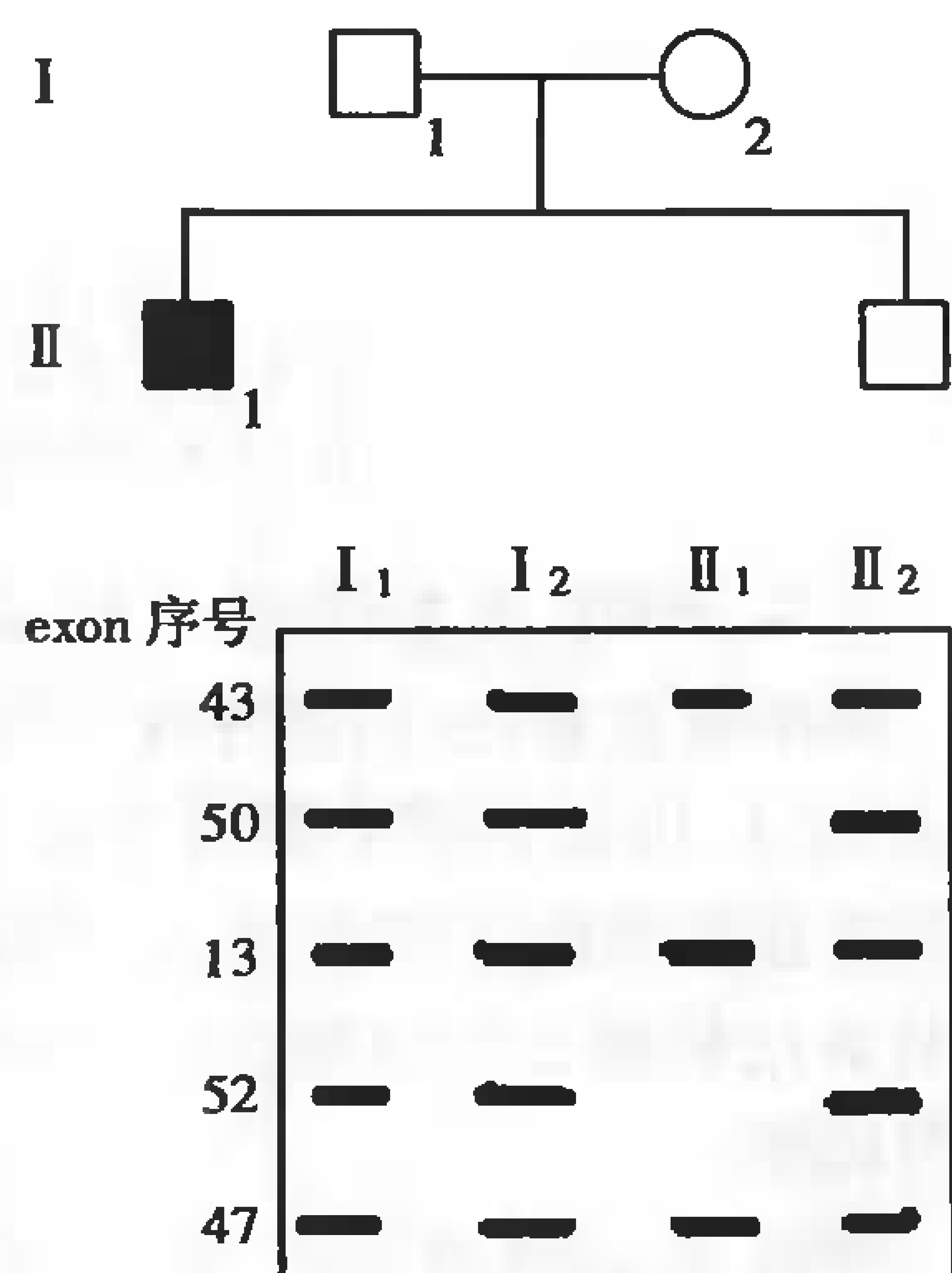


图14-6　DMD家系的多重PCR诊断

2.SSCP - PCR技术　即PCR - 单链构象多态性技术,这是一种基于单链DNA构象差别的快速、敏感、有效的检测基因点突变的方法。基本原理是对已知有基因点突变的遗传病,在其突变位点附近设计引物进行PCR扩增,然后取扩增产物1 μl,在96%的甲酰胺($V:V=1:99$)中变性,然后在不含变性剂的中性聚丙烯酰胺凝胶中电泳,点突变所引起的DNA构象差异将

表现为电泳带位置的差异,据此可做出诊断。

SSCP－PCR 法检测点突变或多态性具有快速、灵敏的优点,但如欲阐明突变的性质,则需做序列分析。

第三节　遗传病的治疗

遗传病通常被认为是一类无法治疗的疾病,然而,大多数遗传病是可以治疗的,少数还可以治愈。随着基因治疗逐步进入临床,遗传病的治疗有了更光明的前景。

一、手术治疗

遗传病造成的器质性病变可进行手术,矫正畸形。例如,先天性心脏畸形可手术矫正,唇裂和(或)腭裂可进行修补,多指(趾)症的多指(趾)可切除,遗传性球形红细胞增多症患者功能亢进的脾可手术切除等。

患者受损的组织或器官可施行组织、器官移植。例如,对胰岛素依赖性糖尿病患者施行胰岛细胞移植,对重型地中海贫血及某些免疫缺陷症患者施行骨髓移植术,对遗传性角膜萎缩症患者施行角膜移植术,对多囊肾患者施行肾移植等。移植肝可提供正常酶源,故肝移植可治疗粘多糖病、α_1 抗胰蛋白酶缺乏症等,这种移植又称为酶移植(enzyme transplantation)。

二、药物治疗

遗传病发展到各种症状出现时,机体器官已造成一定损害,此时药物治疗主要是对症治疗,以缓解病情、减少痛苦。治疗原则是补其所缺、去其所余和激素替代。

(一)补其所缺

即根据遗传病病因,有针对性地补充患者所缺乏的某些成分。分子病及先天性代谢病是蛋白质或酶的缺乏引起,故补充缺乏的蛋白质、酶或它们的终产物,常可收效,但这种补充一般是终生性的。例如,给甲型血友病患者输注抗血友病球蛋白可增强凝血功能,给无丙种球蛋白血症患者补充免疫球蛋白可增强其免疫功能等。

(二)去其所余

由于酶促反应障碍,体内贮积过多"毒物",可使用各种理化方法将过多的"毒物"排出或抑制其生成。可用螯合剂、促排泄剂、代谢抑制剂、平衡清除法、换血或血浆过滤等方法减少体内多余的毒物,从而减轻症状、缓解病情。

肝豆状核变性(Wilson 病)是一种酮代谢障碍的 AR 病,患者细胞内铜离子蓄积造成肝损害、肝肿大、肝硬化,脑基底神经节变性及肾功能损害等临床表现。使用促排泄剂或螯合剂可排出患者体内多余的铜。D－青霉胺是一种螯合剂,可与铜离子结合,加速清除体内的铜离子,对 D－青霉胺过敏或耐药者可采用另一种有效的铜离子螯合剂二盐酸三乙烯四胺。

(三)激素替代

给 Turner 综合症患者补充雌激素可使患者的性器官及第二性征得到发育,也可改善

患者的体格发育;给垂体性侏儒症患者补充生长激素可促进生长发育;给家族性甲状腺肿患者补充甲状腺制剂可增进甲状腺的功能;给糖尿病患者注射胰岛素可使糖尿病症状得到改善。

三、饮食疗法

饮食疗法的治疗原则是禁其所忌,即针对因酶缺乏造成的代谢过程紊乱、底物或中间产物堆积的情况,制定特殊的食谱或配以药物,以控制底物摄入,降低代谢产物的堆积。

饮食疗法成功的先例是苯丙酮尿症的治疗。1953 年,Bickle 等首次用低苯丙氨酸饮食疗法治疗苯丙酮尿症患者,治疗后患者体内苯丙氨酸明显减少、症状减轻、病情缓解。我国北京、天津、上海等地给 1 岁以内的该病患儿食用低苯丙氨酸奶粉也已获得较好的疗效。

再如,半乳糖血症患儿出生后禁饮乳汁,不仅脑功能可发育正常,而且可避免肝损害、白内障等。目前,针对不同代谢病设计和生产了 100 多种奶粉或食谱,可供各种代谢病饮食疗法选用。饮食疗法应用越早,治疗效果越好。

除减少所忌食物的摄入外,减少所忌食物的吸收也可减轻症状,这是饮食疗法的另一途径。例如,苯丙酮尿症患者常规进食后,口服含苯丙氨酸氨基水解酶(phenylalanine ammonialyase)的胶囊,胶囊在肠内释放出的酶能将食物消化产生的苯丙氨酸转化成苯丙烯酸,从而减少苯丙氨酸的吸收;又如,家族性高脂血症Ⅱ型患者服用糠麸,可减少肠内胆固醇的吸收。

四、基因治疗

基因治疗(gene therapy)是指运用 DNA 重组技术,设法修复患者细胞内有缺陷的基因,改变患者细胞的基因表达,使细胞恢复正常功能,从而达到治疗遗传病的目的。

(一)基因治疗的策略和途径

1.基因治疗的策略 根据是否针对患者细胞的致病基因而采取治疗措施,将基因治疗的策略分为直接和间接两大类。

直接策略是针对致病基因的。包括:①基因矫正(gene correction)和基因替换(gene replacement),即矫正缺陷基因的异常序列,精确地原位修复缺陷基因,或定点导入外源正常基因代替有缺陷的基因,对靶细胞的基因组无任何改变。这是理想的基因治疗方式,但由于技术原因,目前尚难实施。②基因添加(gene augmentation),又称补偿性基因治疗,是指不去除异常基因,将有功能的正常基因导入病变细胞,发生非定点整合,表达正常产物以补偿缺陷基因的功能,或使基因的原有功能得以加强。此法难度较小,是目前基因治疗中常用的方式。③基因失活(gene inactivation),又称反义基因治疗,是指采用反义技术、反基因技术或基因敲除技术等,阻断基因的表达。反义技术是将反义 RNA、核酶或反义核酸的表达质粒等导入细胞后,与特定 mRNA 结合,并使其灭活(核酶可切割 mRNA 分子),从而在转录和翻译前水平阻断基因的表达。反基因技术是将设计的寡脱氧核苷酸或肽核酸(一种以多肽骨架取代糖 - 磷酸骨架的 DNA 类似物)导入靶细胞,使寡脱氧核苷酸或肽核酸与靶基因的 DNA 双螺旋分子形成 3 股螺旋,从 DNA 水平阻断或调节基因转录。以肿

瘤细胞中过度表达的癌基因作为靶基因进行此类基因治疗，是肿瘤基因治疗的方向。

间接策略不是针对致病基因，而是导入与靶基因无直接联系的治疗基因，此类策略常用于肿瘤的基因治疗。包括：①免疫性基因治疗，即导入能使机体产生抗肿瘤或抗病毒免疫力的基因。如导入干扰素、肿瘤坏死因子、白介素－2等细胞因子的基因，以增强抗肿瘤效应。②化疗保护性基因治疗，指将编码抗细胞毒性药物蛋白基因导入人体细胞，以提高机体耐受肿瘤化疗药物的能力。如将多药抗性(multiple drug resistance, MDR)基因导入骨髓造血干细胞，减少骨髓受抑制的程度，以加大化疗剂量，提高化疗效果。③自杀基因疗法，是指将来源于病毒或细菌的基因导入肿瘤细胞，该基因产生的酶可催化无毒性的药物前体转变为细胞毒性物质，从而杀死肿瘤细胞。由于携带该基因的受体细胞本身也被杀死，所以这类基因称为"自杀基因"。例如，大肠杆菌胞嘧啶脱氨酶基因导入肿瘤细胞后，可将5－氟胞嘧啶(5－FC)转化为5－氟尿嘧啶(5－FU)，发挥细胞毒性作用。

2.基因治疗的途径　基因治疗的途径有二：一是生殖细胞基因治疗(germ cell gene therapy)；二是体细胞基因治疗(somatic cell gene therapy)。前者是将正常基因导入有遗传缺陷的生殖细胞，使其发育成正常个体，这是根治遗传病的理想方法，可使有害基因不再在人群中散布，但因技术困难及伦理学问题，目前多不考虑这种基因治疗途径。目前的基因治疗途径是后者，即将正常基因导入体细胞使之表达基因产物，以达到治疗目的。

(二)基因治疗的方法

基因治疗的基本过程是应用细胞与分子生物学技术选择并制备目的基因，然后以一定的方式将其导入患者体内，并使该基因有效表达。根据基因导入的方式，基因治疗分为直接体内(*in vivo*)疗法和间接体内(*ex vivo*)疗法。前者是指将目的基因直接导入体内有关的组织或器官，使其进入相应的细胞并进行表达。这是最有前途的方式，但由于技术的限制，目前基因转移和表达的效率还较低。后者是在体外将目的基因导入合适的靶细胞，经过筛选并增殖，然后将细胞回输给患者，使该基因在患者体内有效地表达相应产物，这是目前常用的方式，类似于自体细胞移植。以下以间接体内疗法为例，简述基因治疗的方法。

1.目的基因的选择与制备　目的基因分两类：一是与致病基因相对应的有功能的正常基因，如重症联合免疫缺陷综合征(severe combined immunodeficiency, SCID)是由于腺苷酸脱氨酶(adenosine deaminase, ADA)基因缺陷所致，对该病基因治疗时就选择 *ADA* 基因。二是与致病基因无关，但有治疗作用的基因，如肿瘤治疗中选用的细胞因子基因、自杀基因等。获取目的基因的方法有多种，包括基因克隆、人工合成、PCR扩增等。

2.靶细胞的选择　目前，有多种类型的体细胞可做为基因治疗的靶细胞，包括淋巴细胞、造血干细胞、成纤维细胞、肌细胞、肝细胞、内皮细胞及肿瘤细胞等。靶细胞的选择，根据疾病的特点、基因治疗的策略、目的基因及其转移方式等因素来确定。可以选择病变细胞，也可选择正常细胞。选择原则是：①便于从体内取出和回输；②易于在体外培养与增殖；③便于基因的高效转移；④能够持续表达目的基因。

3.基因转移　如何将外源基因高效地导入特异的靶细胞并稳定表达，这是基因治疗的关键环节。基因转移的方法可分为两大类：①病毒法，主要通过携带有外源基因的病毒载体感染靶细胞来实现基因的转移。能够用作载体的病毒有SV40、牛乳头状瘤病毒、单纯疱疹病毒Ⅰ型、巨细胞病毒、腺病毒和逆转录病毒等。其中较常用的是逆转录病毒和腺

病毒。②非病毒法，即通过物理、化学的方法或受体介导的内吞作用等将外源基因导入细胞内。物理方法包括显微注射、电穿孔、微粒轰击(基因枪技术)及DNA直接注射等；化学方法包括磷酸钙沉淀法、DEAE－葡聚糖法、脂质体融合法等。

4.基因转染细胞的筛选与鉴定　目前基因转移的效率总的来说较低，所以有必要将基因转染的细胞筛选出来，并鉴定该细胞中外源基因的表达状况。例如，可应用标记基因作为探针进行分子杂交的方法筛选，用Northern印迹杂交法来检测RNA的表达，或测定表达产物即蛋白质的含量鉴定细胞中外源基因的表达状况。

5.回输体内　将稳定表达外源基因的细胞经培养、扩增后，以合适的方式，如静脉注射、肌内注射、皮下注射、滴鼻等，回输体内以发挥治疗作用。如将基因修饰的淋巴细胞以静脉注射的方式回输到血液中，将皮肤成纤维细胞以细胞胶原悬液方式注射至患者皮下组织，采用自体骨髓移植的方法输入造血细胞，通过导管技术将血管内皮细胞定位输入血管等。

(三)基因治疗的应用

自从1990年9月14日，Blease等首次对ADA缺乏的SCID患者进行基因治疗获得成功以来，基因治疗的研究迅猛发展，研究范围从单基因遗传病扩展到多基因遗传病，从传统的遗传性疾病扩展到肿瘤、心血管疾病、感染性疾病以及神经系统疾病等，许多治疗方案已从实验室研究过渡到临床应用。据2002年的资料，全世界已有600多个基因治疗临床试验方案，受试患者近4 000例。基因治疗这一具有巨大潜能的新治疗措施，虽然目前尚处于研发阶段，但近年来已取得了一些令人鼓舞的疗效报道。

在遗传病基因治疗方面，由于单基因遗传病的发病机制较为明确，所以单基因遗传病的基因治疗率先取得了一些突破性进展，为其他疾病的基因治疗奠定了基础。至今已有30多种遗传病被列为基因治疗的主要对象，其中*ADA*基因缺陷的SCID、囊性纤维化跨膜传导因子基因缺乏所致的囊性纤维化、低密度脂蛋白受体基因缺陷所致的家族性高胆固醇血症、凝血因子Ⅸ缺陷引起的乙型血友病、葡萄糖脑苷酯酶基因缺乏引起的Gaucher病等疾病的基因治疗研究已获准进入临床试验阶段，并已取得了不同程度的疗效。其中乙型血友病的基因治疗是我国人体基因治疗第1个成功的例子。

在肿瘤基因治疗方面，目前采用的治疗基因非常广泛，既有体内缺陷的基因，但更多的是体内原本不表达、低表达甚至根本不存在的基因，如*MDR*基因。外源基因导入的受体细胞可以是肿瘤细胞，也可以是免疫细胞。治疗策略，除上述间接策略外，还可通过导入特定反义核酸抑制原癌基因的过度表达，或导入肿瘤抑制基因*p*53基因、*RB*基因及肿瘤转移抑制基因*nm*23等抑制肿瘤的发生、发展和转移。目前，肿瘤的基因治疗在体外已取得显著效果。

在心血管疾病的基因治疗中，通过将尿激酶原基因、组织型纤溶酶原激活剂基因导入内皮细胞，再经导管定位导入血管，以达到防止血栓形成的作用；采用反义寡核苷酸封闭*c－myc*、*c－myb*、*N－ras*及胰岛素样生长因子受体基因以抑制平滑肌细胞的增殖；采用心房钠尿肽基因治疗高血压等，这些都为心血管疾病的治疗开辟了新的途径。

此外，基因治疗在感染性疾病，如病毒性肝炎、HIV感染引起的艾滋病，在神经系统疾病的治疗研究中也取得了初步成效。

目前基因治疗研究方面需解决的关键问题是:构建高效、靶向性基因转移系统,外源基因表达的调控,发现切实有效的治疗基因及减少外源基因导入对机体带来的不利影响。人们有理由相信,随着关键技术的突破,21世纪的人类基因治疗将逐步成为一种常规治疗方法,造福于人类。

小　结

遗传病的诊断是遗传病的预防和治疗的前提。遗传病的临床诊断是指从病史人手,通过患者叙述的症状、体格检查获得的体征及系谱分析、皮纹分析的结果,对患者所患疾病是否是遗传病做出初步判断。遗传病的实验诊断包括细胞遗传学检查、生化检查和基因诊断。细胞遗传学检查,即染色体检查和性染色质检查;生化检查,即蛋白质和酶的检测、代谢产物检测;基因诊断是利用DNA重组技术直接从DNA水平上检测人类遗传病的基因缺陷,目前的基因诊断主要是以先证者为线索进行产前诊断。

遗传病的治疗包括手术治疗、药物治疗、饮食疗法和基因治疗。

手术治疗是指对遗传病造成的器质性病变进行手术,矫正畸形,或施行组织、器官移植。药物治疗主要是对症治疗,以缓解病情、减少痛苦;治疗原则是补其所缺、去其所余和激素替代。饮食疗法的治疗原则是禁其所忌,即针对因酶缺乏造成的代谢过程紊乱、底物或中间产物堆积的情况,制定特殊的食谱或配以药物,以控制底物摄人,降低代谢产物的堆积。

基因治疗是指运用DNA重组技术,设法修复患者细胞内有缺陷的基因,改变患者细胞的基因表达,使细胞恢复正常功能,从而达到治疗遗传病的目的。

根据是否针对患者细胞的致病基因而采取治疗措施,基因治疗的策略分为直接和间接两大类。直接策略是针对致病基因的,包括基因矫正和基因替换、基因添加和基因失活。间接策略不是针对致病基因,而是导入与靶基因无直接联系的治疗基因,包括:免疫性基因治疗、化疗保护性基因治疗和自杀基因疗法。此类策略常用于肿瘤的基因治疗。基因治疗的途径包括生殖细胞基因治疗和体细胞基因治疗。

基因治疗的方法包括:目的基因的选择与制备、靶细胞的选择、基因转移、基因转染细胞的筛选与鉴定、稳定表达外源基因的细胞回输体内。

（贡昌春　蔡绍京）

第十五章　遗传病的预防

遗传病是遗传物质改变引起的疾病。目前,大多数遗传病很难治愈,遗传病一旦发生,将可能困扰患者终生,对患者及其家庭造成严重影响。因此,遗传病的预防显得尤为重要。我国于1990年在北京、上海、长沙三地建立了3个国家级遗传医学中心,有计划地指导各地区开展遗传病的预防工作。遗传病的预防工作包括遗传病的普查和登记、遗传咨询、产前诊断、遗传筛查等,其中遗传咨询和产前诊断是防止严重遗传病胎儿出生的重要措施。

第一节　遗传病的抽样调查与登记

要预防遗传病,并控制其在群体中的流行,首先要对某一地区进行选点抽样调查,以了解该地区所存在的遗传病种类、发病率、遗传方式和遗传异质性等的基本情况,估算各种遗传病的基因频率和携带者频率,并对遗传病患者进行登记,以便进一步观察研究。

抽样调查的选点要有代表性,应包括城市、农村、山区等不同区域的人群,以避免人为的选点误差。抽样应包括该地区人口的0.1%~1%,至少包括10万人口,受查率要求在95%以上。普查队伍中需要包括临床各专科的医生及医学遗传学专业人员。抽样调查的遗传病要有明确的诊断标准。

通过抽样调查,对所发现的危害严重的遗传病进行登记,登记内容应包括:①病史,包括姓名、性别、民族、年龄、身高、体重、籍贯、职业、何时迁入本地、所患疾病名称、开始发病年龄、病情演变。②发育史,包括父母是否近亲婚配、患者出生时父母年龄、分娩方式、胎次或产次,母亲妊娠早期有否用药、感染或有放射线接触史,患者有无发育障碍等。③婚姻和生育史,包括配偶姓名、性别、年龄、身高、体重、籍贯、职业、何时迁入本地、与本人是否近亲、健康状况、结婚年龄、是否有自发流产、死产、早产、人工流产等。④家族史,父母年龄、籍贯、职业、何时迁入本地、所患疾病名称、病情进展情况等;同胞数、患者数、发病年龄、病情演变、婚姻状况、生育和健康状况等。其他二级亲属和三级亲属人数、居住地区、健康状况等。

第二节　遗传咨询

遗传咨询(genetic counseling)是由医学遗传学专业人员或遗传咨询医师,应用医学和遗传学基本原理,对咨询者提出的家庭中遗传病的发病原因、遗传方式、诊断、治疗和预后,以及患者同胞、子女再患此病的风险等问题进行交谈和讨论,并就咨询者提出的婚育等问题提出可供咨询者选择的建议或具体指导措施的过程。

咨询医师在遗传咨询过程中,对咨询者应做到"亲切、畅言、守密",有同情心和责任

心，以取得咨询者及其家属的信任与合作，使其能够主动详尽地提供病史和家系资料。要避免使用有损咨询者自尊或刺激性语言来形容患者特征。咨询医师按照遗传病类型和遗传方式估计的遗传病的再发危险，只是发病的概率，不能也不应该作出肯定或否定的保证。

一、遗传咨询的对象

一个人在婚姻或生育方面上遇到问题，意识到可能存在面临遗传病的潜在风险，或者本人或其子女已患遗传病，都应该进行遗传咨询。通常需要进行遗传咨询的情况包括：①夫妇之一患有遗传病或有遗传病家族史；②生育过有遗传病或先天畸形的患儿；③夫妇一方已知或可能是遗传病致病基因的携带者、或染色体平衡易位携带者；④具有不明原因的不孕不育史、复发性流产史和死胎史；⑤35 岁以上的高龄孕妇；⑥近亲婚配的夫妇；⑦有环境有害因素接触史或孕期用药史。

二、遗传咨询的时机

一个人一旦认识到自己可能面临某种遗传病的风险，应该立即进行遗传咨询。理想的初次咨询时间应是在婚前或孕前。有些遗传病患者可在婚前或孕前通过遗传咨询被诊断或做为携带者被检出，获得婚姻或生育指导。如有神经管缺陷生育史的咨询者通过咨询，在孕前开始补充叶酸，可以降低先天性神经管畸形的发生率。已生出某种遗传病患儿的家庭，应尽早遗传咨询，以免失去适当的治疗时机。例如，苯丙酮尿症患儿如果在半岁以内得到确诊而开始治疗，均能收到治疗效果，治疗开始得愈早，患儿智力受损愈轻；如果 1 岁以后才被诊断，则很难治疗，且治疗效果差。对已出生遗传病患儿的夫妇，即使不打算再要孩子，为了解患儿的治疗、预防措施以及将来结婚和生育所面临的问题，仍应该进行遗传咨询。

三、遗传咨询的步骤

虽然随着遗传咨询对象的不同，遗传咨询的步骤有所不同，但各类遗传咨询有其共性的一面，以下以遗传病患者的遗传咨询为例，介绍遗传咨询的一般步骤。

（一）遗传病的诊断

遗传病患者或疑患遗传病的咨询者前来咨询时，作为遗传咨询医师（以下简称咨询医师），首先要对咨询者所患疾病有明确的诊断，如果咨询者所患疾病不是遗传病，就没有遗传咨询的必要。通过询问病史、体格检查、系谱分析，必要时进行细胞遗传学检查或分子生物学检查，可做出是否患有遗传病及相应的遗传方式的诊断（详见第十四章　遗传病的诊断与治疗）。

对于咨询者叙述的病史，系谱分析、体格检查及各项实验室检查结果，咨询医师都要认真登记、填写遗传咨询病历，并妥为保存，以备以后咨询用。

（二）估计再发风险

遗传病诊断确立后，应对再发风险进行估计。再发风险通常用百分率（%）或比例表示。由于公众对医学遗传学的术语不熟悉，所以咨询医师应将再发风险的含义向咨询者

解释清楚。例如,一对已生过 1 个患 AR 病小孩的夫妇,当了解 AR 病的再发风险是 1/4 后会问,是否他们后面的 3 次妊娠都会生下健康的孩子? 这时咨询医师要告诉他们,AR 病再发风险 1/4 是指他们每次妊娠都有 1/4 可能性会生出有病的孩子(详见下节)。

(三)提出对策和措施

在对遗传病做出诊断和再发风险估计后,咨询者所关注的重点是疾病的治疗以及婚姻、生育问题。咨询医师应针对咨询者的问题,与咨询者共同商讨对策,并提出应对措施。值得说明的是,咨询医师在工作中必须贯彻非指令性遗传咨询(non - directive genetic counseling)原则,只提出可供咨询者选择的若干方案,并陈述各种方案的利弊,让咨询者选择,咨询医师不应代替咨询者做决定。

1.对某些可以矫正的生殖器畸形,应待畸形矫正后再结婚。

2.近亲婚配增加 AR 病的发病风险,如果近亲之间正在恋爱或有婚约,应终止恋爱关系或解除婚约。

3.有些遗传病病损严重,可致残或致死,再发风险高于 10%,且不能做产前诊断,对婚后有可能出生这类遗传病患儿的夫妇,应劝其婚后不生育。

4.在先证者所患遗传病较严重,且难于治疗,再发风险高,但患儿父母迫切希望有一个健康孩子的情况下,可运用产前诊断技术选择生育。

5.在已知妻子为 XR 病致病基因携带者、丈夫正常的情况下,可选择生女婴;如为男胎也应进行产前诊断,查明胎儿的基因型,以便决定是否继续妊娠。

6.如丈夫有遗传病不宜生育,可采用供精者精子进行人工授精或采用体外授精 - 胚胎移植技术辅助生殖。

7.当先证者所患遗传病的病损不太严重而且又只有中度再发风险(4% ~ 6%)时,可以冒风险再次生育。例如,一对夫妇已生出了 1 个双侧唇裂的女孩,如果冒险再次生育,将有 4%左右的再发风险,而且其病情严重程度可能有所变化。

8.对目前可以进行植入前诊断的某些遗传病,如囊性纤维变性、脆性 X 综合征、假性肥大性肌营养不良症、染色体数目异常等,可以有目的地选择健康的受精卵植入,以达到预防遗传病的目的。

第三节 遗传病再发风险的估计

再发风险(率)是指家庭中有遗传病患者时,患者亲属再患同样疾病的可能性。再发风险的估计是遗传咨询门诊有别于一般医疗门诊的主要特点。

一、单基因遗传病再发风险的估计

根据亲代基因型是否确定,单基因遗传病再发风险的估计分为 2 种情况。

(一)亲代基因型确定

如果双亲的基因型通过双亲本人或家庭其他成员的患病情况可准确估计,则根据遗传病的不同遗传方式可以推算子女的再发风险。例如,当双亲之一患 AD 病时,此时双亲的基因型可以确定,即 $Aa \times aa$。在外显完全的情况下,子女的再发风险为 1/2;在外显不

完全时,则子女的再发风险为 1/2 乘以外显率。各种遗传方式的单基因遗传病各主要婚配形式下的再发风险在单基因遗传病一章中都已做了分析,此处不再赘述。

(二)亲代基因型不确定

当亲代基因型不确定时,可根据家庭成员的发病情况依据孟德尔定律推算出具有某种基因型的概率,进而对后代的发病风险做出估计,此为概率推算法。如果双亲基因型不确定,在依据孟德尔定律对某个体是否为某种基因型做概率推定的同时,还有其他可利用的信息,如已出生正常子女数、个体年龄、疾病的外显率等,则可运用 Bayes 逆概率定理计算再发风险,此为 Bayes 法。

1.概率推算法 例如,一女性的父母表型正常,但有患红绿色盲(XR)的弟弟,据此可估计该女性是红绿色盲基因携带者的概率为 1/2;如她是携带者,则她与正常男性婚后所生儿子的再发风险为 1/2。但她是携带者的概率只是 1/2,所以,她生儿子的再发风险为 $1/2 \times 1/2 = 1/4$。

再如,在图 15-1 的家系中,先证者 $Ⅲ_1$ 患有 AR 病,他的弟弟 $Ⅲ_2$ 与表妹 $Ⅲ_3$ 结婚,问所生子女发病风险如何?

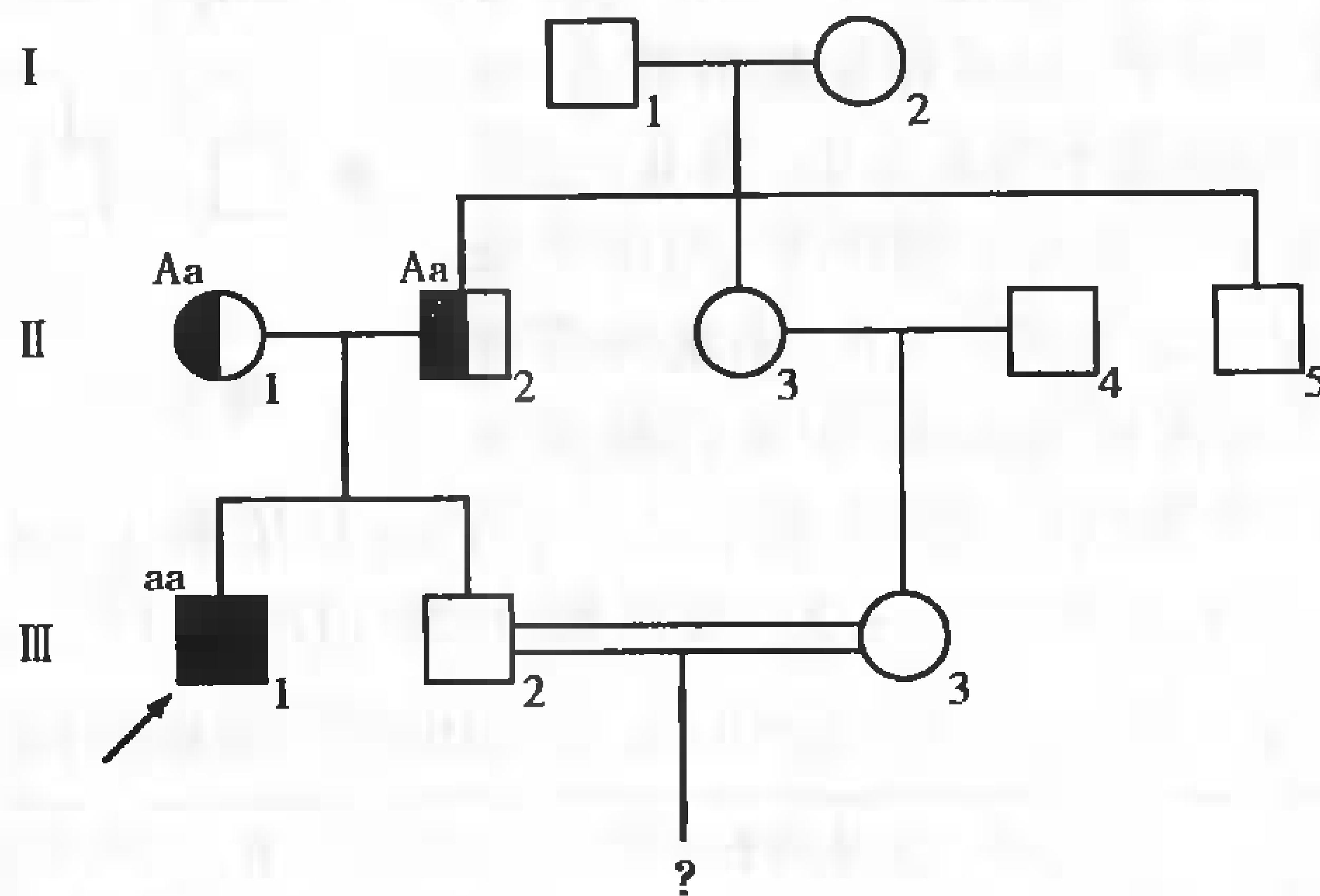


图 15-1　AR 病可能携带者近亲婚配所生子代发病风险的估计

先估计 $Ⅲ_2$ 为携带者(Aa)的概率,因为 AR 病患者表型正常的同胞是携带者的概率为 2/3,所以 $Ⅲ_2$ 为携带者的概率是 2/3。再估计 $Ⅲ_3$ 是携带者的概率,由于 $Ⅲ_1$ 的父亲是肯定携带者(Aa),而 $Ⅱ_2$ 和 $Ⅲ_3$ 为二级亲属,他们之间有 1/4 的基因相同。$Ⅱ_2$ 带有致病基因,则 $Ⅲ_3$ 也带有这一基因(携带者)的概率为 1/4。$Ⅲ_2$ 和 $Ⅲ_3$ 同是携带者的概率为 $2/3 \times 1/4 = 1/6$,他们子女的发病风险则为 $2/3 \times 1/4 \times 1/4 = 1/24$。

2. Bayes 法　Bayes 法的原理是根据事件已发生的结果反过来推算形成这种结果的各种前提的概率。在具体计算时需要作 4 种概率的推算。

(1)前概率(prior probability)　根据孟德尔分离定律推算的某成员具有某基因型的理论概率。例如,AD 病患者子女为杂合子的概率为 1/2,此 1/2 即为前概率。

(2)条件概率(conditional probability)　在某种假设条件下出现某种特定情况的概率。例如,假设夫妻双方都为某 AR 病的携带者,他们生出 1 个不罹患 AR 病孩子的概率为

3/4;生出 3 个均不患 AR 病孩子的概率为 $3/4 \times 3/4 \times 3/4 = 27/64$。这里 3/4、27/64 均为条件概率。

(3)联合概率(joint probability)　为某一基因型前提下,前概率和条件概率说明的 2 个事件同时出现的概率,即前概率和条件概率的乘积。

(4)后概率(posterior probability)　是假设特定条件下的联合概率除以所有假设条件下联合概率之总和,也就是联合概率的相对概率。后概率的计算考虑到了特定条件所提供的信息,因此,它要比前概率更为切合实际。

3.运用 Bayes 法估计再发风险　运用 Bayes 法对某个体是否为某种基因型做概率推定时,可利用的信息包括已出生的正常子女数、个体年龄、疾病的外显率等。

(1)根据正常子女数估计亲代是隐性致病基因携带者的概率

例 1　在图 15-2 的 XR 病系谱中,$Ⅲ_2$ 是 XR 致病基因携带者(以下简称携带者)的概率是多少?

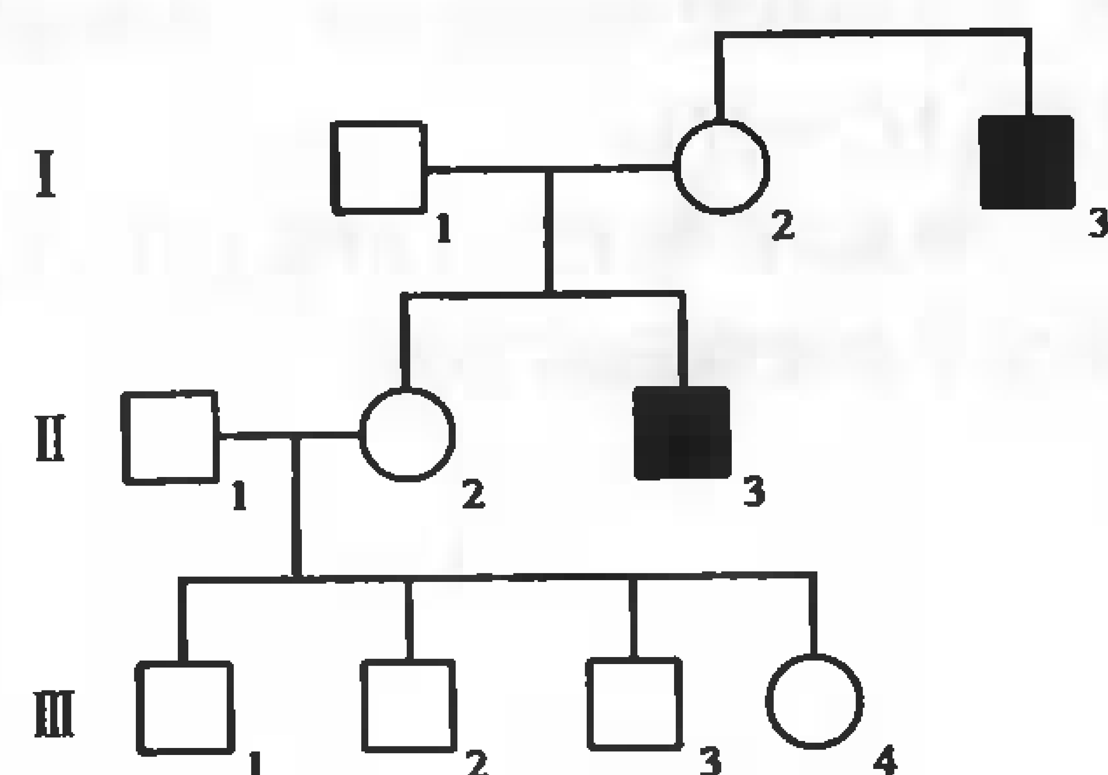


图 15-2　一例 XR 病系谱

分析系谱可知,$Ⅲ_4$ 是 $Ⅱ_2$ 的一级亲属,要知道Ⅲ4 是携带者的概率,必须先知道 $Ⅱ_2$ 是携带者的概率。由于 $Ⅰ_3$、$Ⅱ_3$ 均为患者,故可知 $Ⅰ_2$ 为肯定携带者,而 $Ⅱ_2$ 既可能是携带者,也可能不是携带者,也就是说,2 种情况下的前概率都是 1/2;"当 $Ⅱ_2$ 是携带者时,她能生育 3 个正常男孩的概率"的条件概率为 $1/2 \times 1/2 \times 1/2 = 1/8$;而"当 $Ⅱ_2$ 不是携带者时,她能生育 3 个正常男孩的概率"的条件概率为 1。然后,分别算出 2 种情况下的联合概率:$1/2 \times 1/8 = 1/16$ 和 $1/2 \times 1 = 1/2$。最后,分别算出 2 个后概率:$(1/16)/(1/16 + 1/2) = 1/9$ 和 $(1/2)/(1/16 + 1/2) = 8/9$(表 15-1)。

表 15-1　XR 病家系中 $Ⅱ_2$ 是 XR 基因携带者的概率计算

概　率	$Ⅱ_2$ 是 XR 基因携带者	$Ⅱ_2$ 不是 XR 基因携带者
前概率	1/2	1/2
条件概率	$1/2 \times 1/2 \times 1/2 = 1/8$	1
联合概率	$1/2 \times 1/8 = 1/16$	1/2
后概率	$(1/16)/(1/16 + 1/2) = 1/9$	$(1/2)/(1/16 + 1/2) = 8/9$

由以上分析可知,在已知 $Ⅱ_2$ 有 3 个正常男孩的条件下,$Ⅱ_2$ 是携带者的概率由单纯依靠孟德尔分离定律得到的 1/2 降为 1/9。需要注意的是,不管 $Ⅱ_2$ 基因型如何,其女儿 $Ⅲ_4$ 都不会是 XR 病患者,所以在计算 $Ⅱ_2$ 是携带者的概率时,不需考虑的 $Ⅲ_4$ 的情况。根据 $Ⅱ_2$ 是携带者的概率为 1/9,可得出 $Ⅲ_4$ 是携带者的概率为 $1/9 \times 1/2 = 1/18$。

然而,$Ⅱ_2$ 一旦生育了 1 个患病男孩,那么她就成为肯定携带者,从而推翻了所得出的 Bayes 分析结果。在这种情况下,根据孟德尔分离定律,Ⅲ4 是携带者的概率就恒定为 1/2。

例 2　图 15-3 是 Hurler 综合征的系谱,Hurler 综合征是一种 AR 病。$Ⅲ_1$ 和 $Ⅲ_2$ 为姨表兄妹结婚,他们已有 1 个正常孩子,现前来咨询,问如果再次生育,子代患此病的可能有

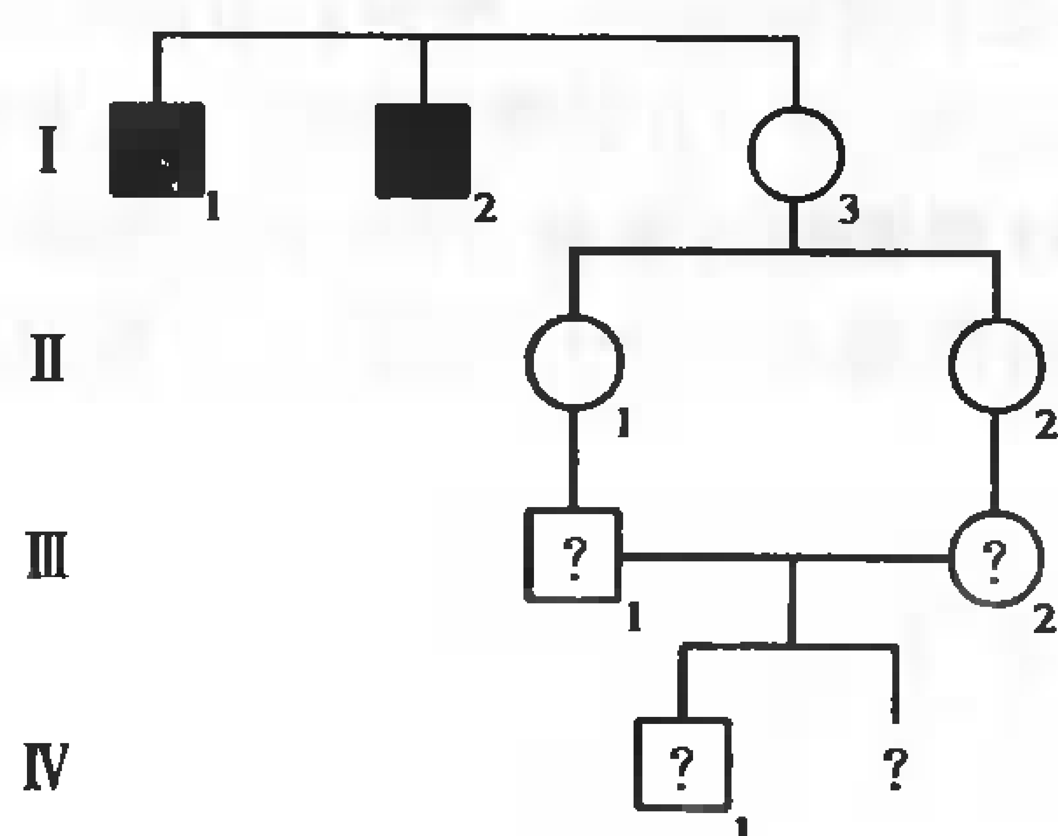


图 15－3　一例 Hurler 综合征系谱

多大？

III_1 和 III_2 后代患病的前提是 2 人都是携带者(Aa)，现已知 I_1 和 I_2 患病，I_3 是携带者的可能性为 2/3，II_1 和 II_2 是携带者的概率均为 1/3，则 III_1 和 III_2 是携带者的概率相同，都为 1/6。两人同时是携带者的概率为 $1/6 \times 1/6 = 1/36$。两人至少 1 人不是携带者的前概率是 $1 - 1/36 = 35/36$。假设两人同时是携带者，则他们生育 1 个正常孩子的可能性为 3/4，这便是两人都是携带者(Aa × Aa)前提下的条件概率。两人至少 1 人不是携带者生下 1 个正常孩子的条件概率为 1。联合概率和后概率的计算见表 15－2。

表 15－2　Hurler 综合征家系中 III_1 和 III_2 同为杂合子的概率计算

概　率	III_1 和 III_2 是致病基因携带者	III_1 和 III_2 不是致病基因携带者
前概率	$1/6 \times 1/6 = 1/36$	$1 - 1/36 = 35/36$
条件概率	3/4	1
联合概率	$1/36 \times 3/4 = 0.02$	$35/36 \times 1 = 0.97$
后概率	$0.02/(0.02 + 0.97) = 0.02$	$0.97/(0.02 + 0.97) = 0.98$

因此两人都是携带者的概率在生育 1 个正常子女后有所下降，由 1/36 降到 1/50 (0.02)。他们第 2 个子女的发病风险为 $0.02 \times 1/4 = 0.005$。

(2)根据外显率估计后代患 AD 病的风险　AD 病的外显率可影响后代的患病风险。例如，已知视网膜母细胞瘤的外显率为 90%，在图 15－4 的视网膜母细胞瘤系谱中，II_1 与正常男性 II_2 婚后所生子女的患病风险如何？

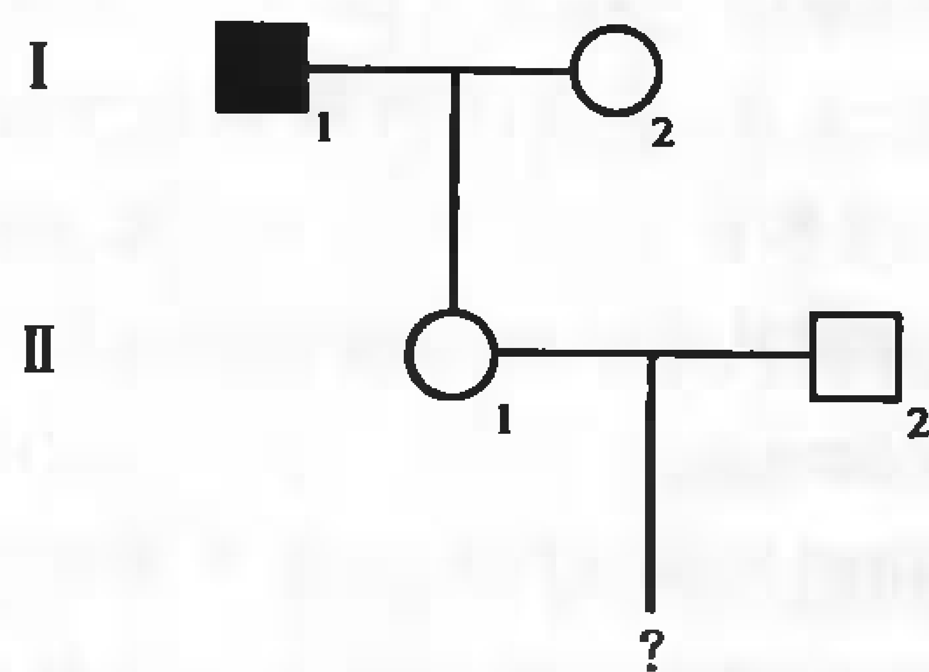


图 15－4　一例视网膜母细胞瘤系谱

按 AD 病的遗传规律，II_1 的基因型是 *Aa* 和 *aa* 的前概率都是 1/2，她的基因型是 *Aa* 但不发病的条件概率为 $1 - 90\% = 10\%$，她的基因型是 *aa* 且不发病的条件概率为 1。据此计算 II_1 的基因型是 *Aa* 的后概率为 9%(表 15－3)。她的子女患病风险则为 $9\% \times 1/2 \times 90\% = 0.0405$。

表 15－3　不完全外显时 II_1 的基因型是 *Aa* 的概率计算

概　率	II_1 为 *Aa*	II_1 为 *aa*
前概率	1/2	1/2
条件概率	$1 - 9/10 = 1/10$	1
联合概率	$1/10 \times 1/2 = 1/$	1/2
后概率	$5\%/(5\% + 50\%) = 9\%$	$50\%/(5\% + 50\%) = 91\%$

(3)根据个体年龄估计迟发 AD 病的再发风险　亨廷顿舞蹈病是一种延迟显性的 AD 病，其外显率随着年龄的增加而增加。在图 15－5 的系谱中，按 AD 的遗传特点，II_2 是患者（*Aa*）的概率为 1/2，II_2 现已 43 岁还未发病。亨廷顿舞蹈病通常在 43 岁前发病的占 64%，未发病的占 36%，即 II_2 是 *Aa* 但 43 岁未发病的条件概率是 36%，据此可计算出联合概率和后概率（表 15－4）。

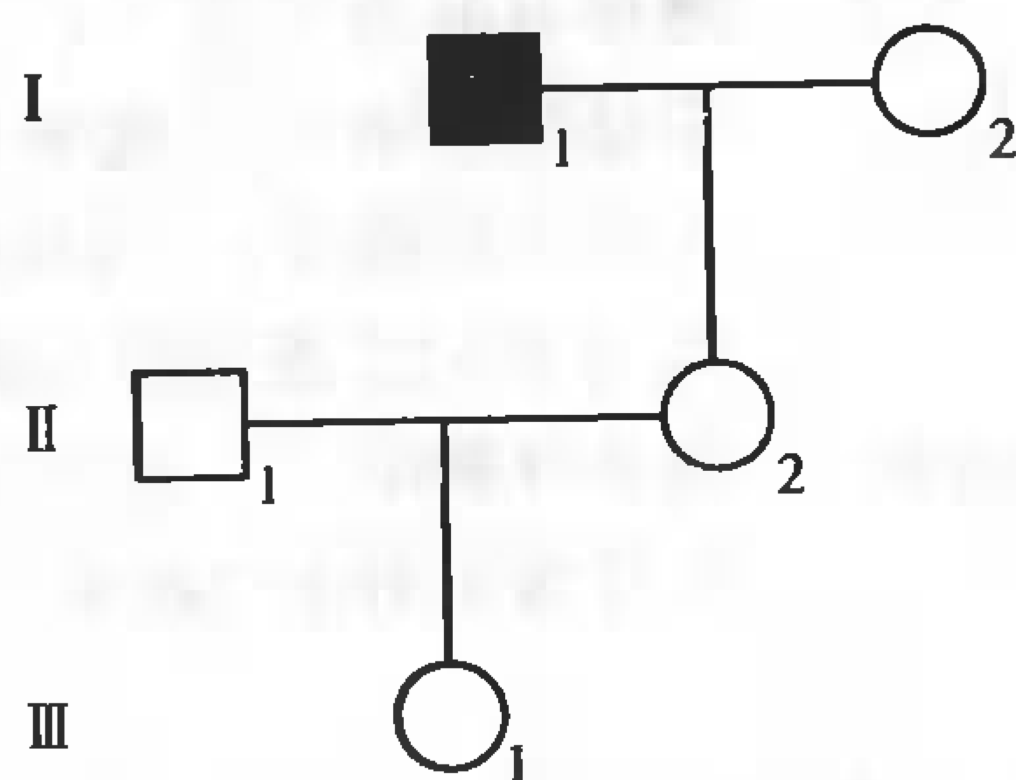


图 15－5　一例亨廷顿舞蹈病系谱

表 15－4　亨廷顿舞蹈病系谱中 II_2 为杂合子的概率计算

概　率	II_2 为 *Aa*	II_2 为 *aa*
前概率	1/2	1/2
条件概率	1－0.64＝0.36	1
联合概率	0.36×1/2＝0.18	1/2
后概率	0.18/(0.18＋0.5)＝0.26	0. 5/(0.18＋0.5)＝0.74

根据以上概率计算可知，由于 II_2 到 43 岁时未发病，她今后患病的概率下降到 0.26，她的女儿 III_1 发病的概率为 0.26×0.5＝0.13。如果再考虑 III_1 到一定年龄没有发病，用 Bayes 法推算后，其发病风险可能还会下降。

再发风险估计是遗传咨询的一项重要工作。一般规定，发病风险高于 10%时为高风险；发病风险介于 1%～10%为中度风险，低于 1%为低风险。在遗传咨询中可根据各种遗传病的发病风险和病损严重程度对咨询者生育进行恰当的指导，或劝阻其生育，或进行适当的产前或症状出现前的诊断措施，或给予疾病的防范指导。

二、多基因遗传病再发风险的估计

多基因遗传病是遗传因素和环境因素共同作用所致，故不能像单基因遗传病那样通过分离律和自由组合率确切算出再发风险率，只能通过群体发病率和家系中受累者的多少加以估计，这种估计概率称为经验危险率（empirical risk）。

多基因遗传病的发病具有下列特点，即亲缘关系越近，再发风险率越大；家系中患病人数越多，再发风险率也越大；该病的遗传率越高，一级亲属的再发风险率也越高。近年来，一些实用的多基因遗传数学模型的相继建立，加上电子计算机的普及，使多基因遗传病再发风险的估计更趋准确。

三、染色体病再发风险的估计

染色体病一般为散发性,其畸变主要发生在亲代生殖细胞的形成过程中,因此再发风险实际上就是经验危险率或称群体发病率。估算染色体病再发风险时,以下几种情况值得重视。

1.父母核型均正常,子女有可能发生染色体数目异常。以后再生子女的再发风险并不高于或稍高于群体发病率。临床上很少见到一个家庭中同时出现2个或2个以上染色体病患者。

2.双亲之一为染色体平衡易位携带者,子女再发风险明显升高,以常见的14/21易位为例,核型为45,XX(XY),der(14;21)(q10;q10)的个体,由于大部分遗传物质并未丢失,这样的个体不会发病而是携带者,其存活的子女中理论上1/3正常,1/3为易位型21三体综合征,1/3为平衡易位携带者。但实际上,14/21易位型21三体综合征的出生率要低于上述理论值,原因可能与流产有关。另外,母亲是平衡易位携带者时,子代发病风险要高于父亲是平衡易位携带者,原因可能在于母亲一般每月只排出1个卵细胞,不像精子那样存在机遇。

3.夫妇之一为染色体病患者,由于染色体病患者多数无生育能力,因此,不存在估计子代再发风险预测;染色体病患者如能生育,子代的发病风险将明显高于群体发病率。

4.母亲年龄的增大,生育染色体病患儿(尤其是三体型)的风险逐渐升高,可能是由于35岁以上的妇女的卵巢功能开始退化,导致卵细胞成熟过程中易于发生染色体不分离所致。

第四节 产前诊断

产前诊断(prenatal diagnosis)又称宫内诊断或出生前诊断,是在遗传咨询的基础上,在胎儿出生之前,对有高风险的孕妇进行特殊检查,了解胎儿在子宫内的发育情况,对胎儿是否患有某种遗传病或先天畸形做出的诊断。如果确认胎儿患有严重遗传病或先天畸形,则终止妊娠。因此,产前诊断是优生的重要措施之一。

目前可以进行产前诊断的遗传病包括:染色体病,特定酶缺陷所致的先天性代谢病,可进行DNA检测的遗传病,多基因遗传的神经管缺陷,有明显形态改变的先天畸形。

根据临床实践,结合目前可进行产前诊断的疾病分类,通常认为有以下情况之一者应进行产前诊断:①孕妇年龄≥35岁;②夫妇一方染色体异常或曾生育过染色体病患儿;③曾生育过先天畸形(特别是神经管缺损)患者;④作为AR病或XR病携带者,或生育过某种单基因遗传病患儿;⑤原因不明的流产、死产、畸胎和新生儿死亡史等;④丈夫孕前、妻子孕早期接触致畸因素,如化学毒物、辐射、病毒感染等。

一、产前诊断技术及应用

产前诊断主要以羊膜穿刺、绒毛取样和胎儿脐血穿刺等技术为手段,获取胎儿的组织,如羊水、绒毛、脐带血等,通过获取的胎儿组织进行细胞遗传学检查、生化分析、基因诊断等,判断胎儿是否患有遗传病。

(一)羊膜穿刺术

羊膜穿刺(amniocentesis)一般在妊娠16～20周进行,此时羊水量多,穿刺成功率高。羊水中胎儿细胞易于培养。穿刺前,先嘱孕妇排空膀胱,然后经B型超声检查,选择穿刺点;穿刺点应避开胎盘,选取胎儿肢体侧或胎儿颈后部较空虚的部位。穿刺时,孕妇取仰卧位,腹部消毒后,用腰穿针自穿刺点垂直刺入;穿刺针刺穿腹壁和子宫壁时有2次突破感,进入羊膜腔有针尖的落空感,拔出针芯后见有清亮羊水流出。穿刺成功后抽取羊水20～30 ml。

获得的羊水可有如下用途:①羊水细胞可进行染色体核型分析,诊断染色体病;②通过测定羊水中各种酶的含量,诊断先天性代谢病,如羊水中甲胎蛋白和乙酰胆碱酯酶的含量升高,提示开放性神经管畸形;③检测Y染色体上SRY基因,预测胎儿性别,有助于XR病的诊断;④提取胎儿细胞的DNA进行基因诊断。

(二)绒毛取样法

绒毛取样(chorionic villus sampling,CVS)一般在妊娠6～9周进行。术前应用B型超声检查确定子宫位置。具体操作方法为:孕妇排空膀胱后,取膀胱截石位,消毒外阴、阴道及宫颈。在B超引导下将一直径2 mm带芯塑料管经宫颈管沿宫腔方向贴近宫壁缓慢进入宫腔,达到绒毛处后,退出针芯,用含细胞培养液的5ml注射器抽取少量绒毛组织(15～40 mg)即可。

绒毛细胞是胎儿细胞的组成部分,通过绒毛细胞的染色体核型分析可诊断胎儿的染色体病;根据核型显示的胎儿性别,有助于早期发现X连锁遗传病;绒毛细胞富含各种酶,可经生化方法检测先天性代谢病;绒毛细胞的DNA分析,可对胎儿进行基因诊断。

(三)脐带穿刺术

脐带穿刺术(cordocentesis)又称经皮脐静脉穿刺取血术,是指在B型超声引导下经孕妇腹壁穿刺抽取胎儿脐静脉血的方法。此法能获得纯胎儿血标本用于产前诊断,一般在妊娠18周后进行。脐静脉直径应在4 mm以上。术前准备与羊膜穿刺术相同。首先用B型超声选定穿刺点,一般选择脐蒂处,此处较为固定,便于穿刺。穿刺在B型超声引导下进行,当确定针尖进入脐静脉后,先试抽0.2～0.3 ml,证实为脐血后,再根据需要取血2～8 ml。

胎儿血的用途十分广泛,可用于各种生化检查,诊断先天性代谢病;胎儿血中的有核细胞可进行染色体分析;提取细胞中DNA可用于基因诊断。

(四)其他产前诊断技术

1.超声检查　超声检查简便无创伤,目前最常用的是B型超声仪。一般认为超声强度在20 mW/cm^{2}以下,持续时间不超过30 min,对孕早期3个月以内的胎儿是安全的。B型超声仪能显示胎儿、胎盘及羊水情况,可诊断各种胎儿畸形;在妊娠18周以后,可以辨认胎儿性别,用于XR病的产前诊断。彩色多普勒超声仪可显示胎儿心脏情况,用于检查胎儿先天性心脏病,还可用于胎儿血流的检查。

2.磁共振检查　妊娠16周后,胎儿四肢长骨、短骨和肋骨已经骨化,可通过磁共振显像。磁共振检查效果显著优于X线检查,并且X线对胎儿有一定危害性,故目前已不用X线对胎儿进行产前诊断。磁共振检查主要用于了解胎儿有无先天性畸形,如无脑儿、脑积水、骨骼畸形、侏儒、多指、短指或缺肢、脊柱裂及胸廓畸形等。

3.孕妇外周血中胎儿细胞或胎儿 DNA 检查 取胚胎绒毛、抽取羊水或脐带血等均为侵袭性方法,可能对孕妇及胎儿造成一定的危害,诸如流产、宫内感染和胎儿损伤等;利用孕妇外周血中胎儿细胞或胎儿 DNA 进行产前诊断是最近几年发展起来的一种非侵袭性产前诊断新技术。研究表明,在孕 5~7 周时即可从孕妇外周血中分离到胎儿细胞或提取出胎儿 DNA;随着妊娠进展,孕妇外周血中胎儿细胞及胎儿 DNA 的量不断增高。目前,此项技术的问题是从母血中得到的胎儿细胞或胎儿 DNA 数量太少,且方法复杂,亦不能有效地排除母体细胞和母体 DNA 的干扰,因此,此项技术还有待进一步研究。

二、常见遗传病的产前诊断

(一)21 三体综合征

1.胎儿细胞染色体核型分析 早期妊娠,通过绒毛取样法获取绒毛细胞;中、晚期妊娠,通过羊膜穿刺术或脐带穿刺术获取羊水细胞或脐带血。通过绒毛细胞的直接制片,或羊水细胞、脐带血淋巴细胞培养,可对胎儿进行染色体核型分析,这是产前 21 三体综合征的确诊方法。

2.B 型超声监测 B 型超声监测是 21 三体综合征的一种筛查手段。在妊娠早期,胎儿颈背皮下组织开始积蓄一些半透明液状物,称为颈背透明物(nuchal translucency,NT),也称颈背水肿、颈背褶。NT 厚度的增加与 21 三体综合征有密切的关系,可以作为 21 三体综合征高危孕妇的指标,通常妊娠 10~13 周时,NT 厚度小于 3 mm,21 三体综合征胎儿 NT 厚度多超过此范围。

3.母体血清生化标志物筛查 孕期胎儿肝脏合成 AFP,21 三体综合征的胎儿常有肝脏或胎盘功能不良,合成 AFP 的能力降低,故母体血清 AFP 常降低。另外,临床研究表明,妊娠中期,怀有 21 三体综合征胎儿的孕妇血清中,游离的 β 型人绒毛膜促性腺激素(free β-HCG)水平高出正常 2 倍以上。因此,这 2 种孕妇血清标志检测有助于筛查出 21 三体综合征阳性胎儿。根据孕妇血清 AFP 和 Free β-HCG 的水平,结合孕妇年龄、孕周、体重等资料,利用计算机软件进行分析,可以进一步提高筛查的阳性率。值得注意的是,筛查结果阳性只说明胎儿患 21 三体综合征的风险升高,应该做进一步的确诊检查。

(二)神经管缺陷

神经管缺陷(neural tube defect,NTD)属多基因遗传病,临床上分为无脑儿、脊柱裂和脑疝 3 种,也可分为开放性 NTD 和闭合性 NTD 2 种。超声检测对无脑儿和脊柱裂的产前诊断结果准确可靠,对开放性 NTD,还可检测孕妇血清 AFP 水平和胎儿羊水乙酰胆碱酯酶(AchE)的含量。

在开放性 NTD,胎儿的 AFP 大量漏出,羊水 AFP 含量显著增高,从而使母体血清中的 AFP 浓度显著升高。通过测定母体血清和胎儿羊水的 AFP 水平,可对开放性 NTD 进行筛查。正常情况下,胎儿脑脊液中 AchE 浓度很高,而血清中浓度很低。当胎儿患开放性 NTD 时,脑脊液中的 AchE 大量漏到血液里,然后通过胎膜渗透到羊水里。因此,可通过检测羊水 AchE 水平对开放性 NTD 进行辅助诊断,确诊率达 99%以上。

(三)甲型血友病

甲型血友病是凝血因子Ⅷ基因突变所致的 XR 病。当孕妇为携带者、丈夫正常时,儿

子为患者、女儿为携带者的概率都是1/2。因此,甲型血友病产前诊断包括风险胎儿性别诊断、男性胎儿基因诊断和女性胎儿携带者基因诊断3个方面。另外,胎儿血中凝血因子Ⅷ抗原的测定也可作为产前诊断的辅助方法。

通过胎儿绒毛细胞或羊水细胞或脐带血淋巴细胞可进行染色体核型分析,做出胎儿性别的诊断;如为男性胎儿需进行甲型血友病基因诊断。由于基因突变存在高度遗传异质性,常表现为家族特异性,因此直接诊断基因突变的方法往往仅限用于某一家族。凝血因子Ⅷ基因存在基因内或基因侧翼RFLP多态位点以及"CA"串联重复多态性位点,所以,通过对这些多态性位点的连锁分析可进行患者和携带者的筛查。此法不足之处是可因基因内重组而使诊断结果错误,因此,如有可能,应首选基因突变的直接诊断法。

根据凝血因子Ⅷ基因突变的类型及其频率,目前国外对突变类型未知的甲型血友病病例进行基因诊断的方案采取3大步骤:首先决定是否有倒位突变;如果倒位突变阴性,第二步检测最常见的内含子13和内含子22的$(CA)_n$ STR;最后进行内含子18/BclⅠ和内含子22/XbaⅠ的RFLP多态性位点的连锁分析。通过这样的检测方案,可使约85%的患病家系得到可靠的诊断。

现以图15-6的家系为例,说明甲型血友病的基因诊断方法。已知甲型血友病基因的正常等位基因内含子18包含BclⅠ的酶切位点,酶切后产生99 bp DNA片段。如果内含子18的BclⅠ酶切位点发生基因突变,酶切后则产生143 bp DNA大片段。

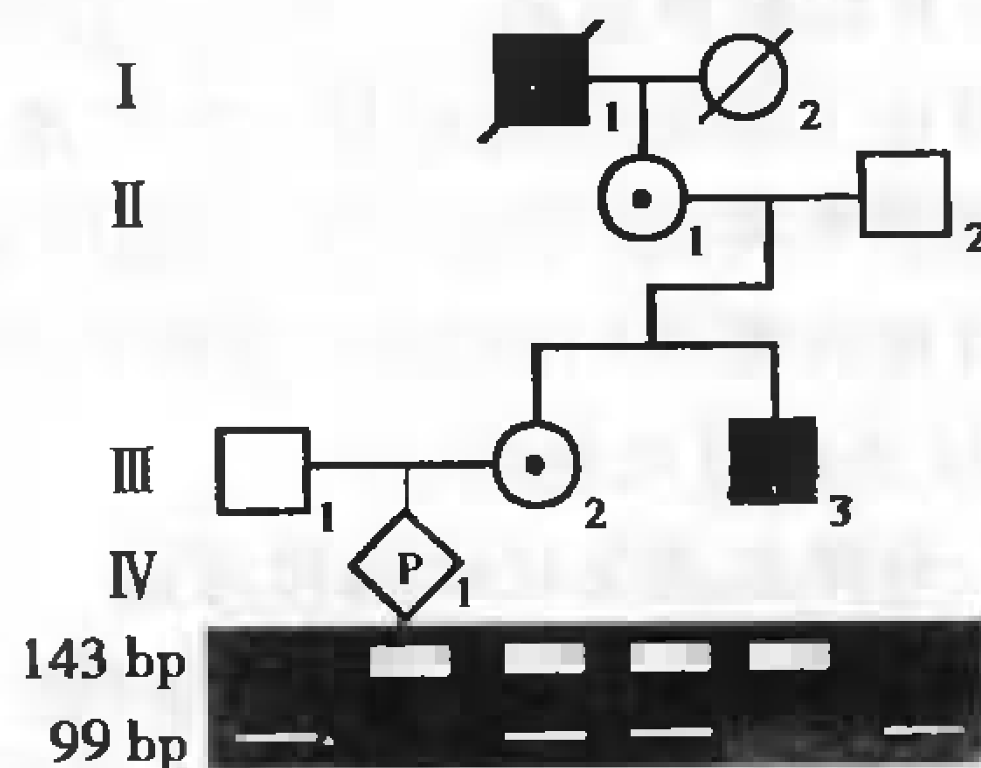


图15-6 甲型血友病产前基因诊断

在该家系中,$Ⅲ_3$是患者,因失去BclⅠ切点,故只有143 bp片段;$Ⅱ_1$的父亲$Ⅰ_1$是患者,$Ⅱ_1$是肯定携带者,2条X染色体上的BclⅠ酶切位点1个失去,1个正常,故有143bp和99 bp 2种片段;根据$Ⅲ_2$出现143 bp和99 bp 2种片段,可知她也为肯定携带者;现$Ⅳ_1$尚未出生,根据其只有143 bp片段,可知他是男性患者。

第五节 婚姻指导及生育指导

对遗传病患者及其亲属进行婚姻和生育指导可减少遗传病患儿的出生,降低群体的致病基因频率,是遗传咨询医师的重要任务之一,此项工作常结合遗传咨询进行。

一、婚姻指导

婚姻指导包括婚前检查和避免近亲结婚,这是保障后代健康,预防遗传病患儿出生的重要措施之一。

婚前检查着重对双方重要脏器及生殖器官进行检查,特别注意生殖器官有无畸形,必要时检查性腺的发育及功能。婚前检查时可进行婚前咨询,了解双方是否为近亲,双方家庭及亲属有无遗传病史。

AD病患者子代患病风险达50%，且病情多较重，一般不宜结婚。隐性遗传病杂合子婚配，是生育重症遗传病患儿的最主要来源。因此，必须劝阻杂合子间婚配。在尚无条件进行杂合子检测时，应尽量避免近亲结婚。例如，群体中苯丙酮尿症的致病基因频率为1/100，杂合子频率为1/50，随机婚配出生纯合子患儿的概率为1/50×1/50×1/40=1/10 000；如为表兄妹婚配，则出生患儿的概率为1/50×1/8×1/4=1/1 600，两种婚配情况相比，出生隐性遗传病患儿的风险相差6倍多。隐性遗传病发病率越低，近亲结婚生育患儿的相对风险较非近亲结婚者越高。据世界卫生组织调查，非近亲婚配所生婴儿的死亡率为24‰，而近亲婚配者则为81‰。因此，避免近亲结婚对预防遗传病相当重要。对遗传病患者或其亲属的其他婚姻方面的指导，参阅本章遗传咨询的相关内容。

二、生育指导

（一）孕前指导

首先要提倡适龄生育，我国古代就有“男虽十六而精通，必三十而娶；女虽十四而天癸（月经）始至，必二十而嫁。皆欲阴阳充实，然交而孕、孕而育、育而子坚壮强寿。”的说法。大多数妇女生育旺盛期是20~30岁，从产科角度看，这个年龄段的生育和分娩对母婴健康是适宜的。在发达国家，25~29岁年龄段为生育高峰期。年龄过小或过大生育，先天畸形的发生率都很高。统计资料显示，与25~29岁生育相比：20岁以下生育，先天畸形的发生率高50%；35岁以上生育，先天愚型发生率高5倍，40岁以上生育则高出10~15倍。40岁以上的高龄孕妇往往伴有较多的妊娠并发症，加之骨盆各韧带的弹性和松弛性下降，易造成分娩困难，导致新生儿产伤和死亡率增加。男子超过35岁，精子的染色体异常增加，精子的基因突变率显著增高。因此，男性也不应过晚生育。

其次，夫妻双方健康的身体及良好的受孕环境也是非常重要的。美国国立精神健康研究所的研究成果表明，男女之情的萌发与脑垂体后叶分泌的加压素和催产素有关。大量的加压素和催产素，加上性激素，构成了有利于生殖的内环境。“情深婴美”说的就是这个道理。另外，家庭环境（住房、饮食等）、气候条件与社会环境（战争、饥荒、动乱等）对受孕均有影响。

（二）孕期指导

胚胎发育的早期（第12周前）是胚胎器官发生的关键时期，此期内任何环境因素的作用均可干扰胚胎器官的正常发育而形成先天畸形。这些环境因素包括生物因素（如风疹病毒、巨细胞病毒、单纯疱疹病毒、弓形虫等）、物理因素（如γ射线和X线等）及化学因素（各种化学试剂、环境污染等）等，孕妇应尽可能避免接触这些影响胚胎发育的环境因素。

孕期特别要注意安全用药，尽可能避免用药或在医生指导下用药，以避免对胎儿的伤害。要注意有意识地进行孕期胎教。胎儿是有意识、能反应的小生命，母亲怀孕期的精神因素、心理状态等都会影响胎儿，故孕妇应保持良好的心境。孕期的饮食既要营养丰富，又要容易吸收，以“清淡”为主；养成良好的生活习惯，不吸烟、不饮酒。

定期进行产前检查：若夫妇均为AR病携带者，应做产前诊断，以确定胎儿是否为患者；在XR病的家系，应做产前诊断确诊胎儿性别；根据产前诊断的结果，决定继续妊娠还是人工流产。

(三)辅助生殖

当夫妇一方患单基因遗传病或染色体病,或夫妇双方是致病基因携带者时,为了避免遗传病患儿的出生,虽然可通过产前诊断确认胎儿是否患有遗传病,然后进行选择性人工流产。但大多数遗传病目前尚无有效的产前诊断方法。如果夫妇一方存在生殖器官异常,或不能产生正常配子,则无法正常生育。在这些情况下,可采取辅助生殖技术,帮助这些夫妇生育正常子女。目前开展的辅助生殖技术主要包括人工授精和体外授精 - 胚胎移植。

1.人工授精(artificial insemination)　是指用人工的方法将精子置入女性生殖道内,使卵子受精并发育。根据精子的来源,人工授精分为夫精人工授精和供精人工授精 2 类。

夫精人工授精(artificial insemination by husband,AIH)适用各种原因所致的丈夫无法正常性交受精者。供精人工授精(artificial insemination by donor, AID)适应于男性不育患者、男性常染色体显性遗传病患者、夫妇同为隐性遗传病携带者以及 Rh 血型不合者等。可以使用健康男性的新鲜精液,也可使用冷冻精液。冷冻精子库的建立推动了人工授精的广泛开展。1983 年,我国湖南医科大学人类生殖工程研究室用冷冻精液进行人工授精获得成功。

2.体外受精与胚胎移植(*in vitro* fertilization & embryo transplantation,IVF - ET)　是指用人工方法(通常使用药物激发排卵)获得卵子,在体外使卵子受精并进行早期胚胎发育,发育至 4 ~ 8 个细胞或囊胚阶段时,将卵裂球或囊胚植入母体子宫内继续发育直到出生。此种技术通常称为"试管婴儿"技术,采用此种类技术出生的婴儿称为"试管婴儿"。

1978 年 7 月 25 日,世界第 1 例试管婴儿在英国诞生;此后,在澳大利亚、美国、法国、瑞典、丹麦、日本等国相继有试管婴儿诞生;1988 年 3 月,我国首例试管婴儿在北京医科大学诞生。

IVF - ET 主要适应证是输卵管因素导致的女性不育,如输卵管闭塞、子宫内膜异位症等。在女性患这些疾病时,精子和卵子无法在体内正常受精。

如果一对夫妇曾出生过严重 AR 病患儿,或妻子为 AD 病患者或染色体相互易位携带者,为避免再次出生患儿,可由健康供卵者提供卵子进行 IVF - ET;完成早期胚胎发育后,再将卵裂球或囊胚植入到妻子的子宫内继续发育。这种辅助生殖技术通常称为"借卵怀胎"。

如果男性患严重的少精症、弱精症及阻塞性无精症等疾病,可采用卵母细胞单精子显微注射(intracytoplasmic sperm injection, ICSI)技术进行体外受精,ICSI 技术是从 IVF - ET 衍生而来。对于有较高风险生出遗传病患儿、而又无有效的产前诊断方法的夫妇,可应用 IVF - ET 衍生的另一项技术——胚胎植入前遗传学诊断(preimplantation genetic diagnosis, PGD)技术对"试管婴儿"进行诊断。PGD 技术包括:①卵裂球活检:在 IVF - ET 胚胎发育至 6 ~ 10 个细胞时,取 1 ~ 2 个卵裂球进行相关诊断后,将正常胚胎移植入母体子宫继续发育;②极体活检:对卵母细胞排出的第一极体和第二极体活检,主要用于卵细胞染色体数目异常的分析;③滋养层细胞活检:在囊胚期对滋养层细胞活检可获得较多细胞(10 ~ 30 个),能进行染色体核型分析。

3.重组 DNA 技术　转基因技术是近年来生命科学中发展最快的领域之一。1980 年,

Gorden 等首次采用显微注射的方法将外源性遗传物质导入小鼠受精卵的原核中，获得了基因重新组合的小鼠模型。目前研究较多的从体外受精及显微操作技术发展而来的精子介导基因转移及核移植技术、基因打靶(gene targeting)技术可将外源 DNA 序列与活细胞内染色体发生同源重组以定点修饰染色体上某一特定基因。

基因重组技术应用于生殖研究领域，有可能对某些异常的基因和染色体进行修饰和替换，从而提高人类的遗传素质。

第六节 遗传筛查

遗传筛查是依据某些指标对某一群体的某种遗传病进行普查，初步筛选出可疑人群，然后再进一步确诊的过程。通过筛查，可较早地发现尚无临床表现的患者或致病遗传物质携带者，并尽早采取预防措施。

一、新生儿筛查

新生儿筛查(neonatal screening)是对已出生的新生儿进行某些遗传病症状前诊断，是出生后预防和治疗某些遗传病的有效措施。选择筛查的病种通常应符合如下条件：①发病率较高；②有致死、致残、致愚的严重后果；③有较准确且经济实用的筛查方法；④筛出的疾病有治疗措施。

目前我国列入新生儿筛查的疾病有苯丙酮尿症、先天性甲状腺功能低下、G6PD 缺乏症(南方)。

新生儿筛查一般是以静脉血或尿为材料。血样的采集在新生儿出生后的 3~4 天进行，从足跟部采血，用滤纸吸全血晾干形成血斑。尿样采集的方法是在新生儿的尿布中夹放滤纸或直接收集 1~2 ml 新鲜尿液。

(一)苯丙酮尿症的筛查

苯丙酮尿症是由于苯丙氨酸羟化酶缺乏所引起的一种先天性代谢病，临床表现严重，并伴有智力发育异常，但如能在新生儿期发现，则可以通过饮食控制等措施防止或减缓症状的出现和发展。目前，国际上常用的筛查方法是 Guthrie 细菌抑制法。方法是将干血片置于含有与苯丙氨酸结构相似的细菌抑制剂β-2-噻嗯丙氨酸的培养板上，在 37 ℃温箱中培养 24 h，观察干血片周围的细菌生长环。正常标本，枯草杆菌受到培养板中抑制剂的作用不能生长或出现较小的生长环。如果血液标本中的苯丙氨酸浓度增高，则细菌的生长不受抑制，培养板上可出现较大的细菌生长环。筛查出的阳性患者均应采取静脉血作苯丙氨酸及酪氨酸测定。明确诊断者，应立即进行饮食治疗，减少苯丙氨酸的摄入。

(二)先天性甲状腺功能低下的筛查

先天性甲状腺功能低下多数是由于甲状腺发育异常所致。该病的发病率高，早期诊断方法简单，治疗疗效明显。采用血斑滤纸的提取液以 ELASA 法测定 TSH 和 T_4 含量。先天性甲状腺功能低下的患儿 TSH 水平升高而 T_4 水平降低。一旦明确诊断，应尽快补充 T_4，以抑制 TSH 的增高。

二、携带者筛查

遗传携带者(genetic carrier)是指表型正常,但带有致病遗传物质(致病基因或畸变的染色体),能遗传给后代并使之患病的个体,包括:①隐性遗传病杂合子(简称杂合子);②显性遗传病未显者;③表型尚正常的迟发外显者;④染色体平衡易位携带者。

携带者筛查是指当某种遗传病在某一群体中有较高发病率,为了预防该病在群体中的发生,采用经济实用、准确可靠的方法在群体中进行筛查的过程。这种筛查主要是对杂合子的筛查,对显性遗传病的未显者及表型尚正常的迟发外显者目前尚无有效的筛查方法。染色体平衡易位携带者的筛查只在相关家系中进行,由于染色体平衡易位携带者生育死胎及染色体病患儿的机会很大,因此,染色体平衡易位携带者的筛查十分重要。

人群中许多隐性遗传病的发病率不高,但杂合子的比例却相当高,所以,杂合子的检出对遗传病的预防尤为重要。对发病率很低的遗传病,一般不做杂合子的群体筛查,仅对患者亲属进行筛查。对发病率高的遗传病,普查携带者效果显著。

例如,我国南方各省的α及β地中海贫血的发病率高达8%~12%,有的省或地区更高,因此,检出夫妇双方同为α或同为β地贫杂合子的机会很多。这时,进行婚姻及生育指导,配合产前诊断,就可从第1胎起防止重症患儿出生,不仅降低了该病的发病率,而且防止了致病基因在群体中播散。再如,在犹太人群中,进行血清氨基己糖苷酶活性检测,可筛查Tay-Sachs病携带者,凡酶活性降低者,可确认为携带者。对携带者进行婚育指导,并对后代中有风险的孕妇实施产前诊断,如确诊胎儿为患者,则终止妊娠。通过此种筛查方法,在过去25年,犹太人Tay-Sachs病的发病率显著下降。

杂合子检测方法的理论根据是基因的剂量效应,即杂合子基因产物的量介于纯合子与正常个体之间,约为正常个体的半量。但由于机体内、外环境因素对基因表达的影响,以及检测方法的不同(直接测定基因产物或测定基因间接产物),常使杂合子测定值偏差,造成判断困难。

杂合子的检测可在不同水平进行。临床水平的观察一般只能提供线索,不能准确检出。例如,某些红细胞酶缺乏症的杂合子有轻度贫血表现,眼白化症女性杂合子的眼底可呈现虎斑状镶嵌色素沉着。细胞水平主要进行染色体检查和组织学观察。例如,通过染色体检查可发现平衡易位携带者和脆性X染色体综合征的女性携带者,通过组织学和细胞学观察,可发现镰形细胞贫血症携带者的红细胞呈轻度镰变。蛋白质水平的检测,包括酶及代谢中间产物的测定,对一些分子病、先天性代谢病杂合子检测正逐渐被基因水平的方法所取代。

随着分子遗传学的发展,不但可以从分子水平即利用DNA或RNA分析技术直接检出杂合子,而且结果更准确,特别是对一些致病基因的性质和异常基因产物还不清楚的遗传病,或用一般生化方法不能准确检测的遗传病。例如,甲型和乙型血友病、苯丙酮尿症等。目前,用基因分析检测杂合子的方法日益增多,并逐步向简化、快速、准确的方向发展,逐步扩大到高危人群的筛查。

小 结

大多数遗传病很难治愈,遗传病一旦发生,将对患者及其家庭造成严重影响。因此,遗传病的预防显得尤为重要。遗传病的预防工作包括遗传病的普查和登记、遗传咨询、产前诊断、遗传筛查等。

遗传咨询就是被咨询者应用医学和遗传学基本原理,对咨询者提出的家庭中遗传病的发病原因、遗传方式、诊断、治疗和预后,以及患者同胞、子女再患此病的风险等问题进行交谈和讨论,并就咨询者提出的婚育等问题提出可供咨询者选择的建议或具体指导措施的过程。

单基因遗传病再病风险的估计分为 2 种情况:亲代基因型确定时,可根据遗传病的不同遗传方式推算再发风险;当亲代基因型不确定时,可应用概率推算法和 Bayes 法估计再发风险。Bayes 法是根据已出生正常子女数、个体年龄、疾病的外显率等信息,运用 Bayes 逆概率定理计算再发风险的方法。

产前诊断是在遗传咨询的基础上,通过特殊的检查方法,对胎儿是否患有某种遗传病及先天畸形做出的诊断。如果确认胎儿患有严重遗传病或先天畸形则终止妊娠。因此,产前诊断是预防遗传病患儿出生的重要措施。产前诊断常用方法有羊膜穿刺术、绒毛取样法和脐带穿刺术等。

对遗传病患者及其亲属进行婚姻和生育指导,必要时采用辅助生殖技术,可减少遗传病患儿的出生,降低群体的致病基因频率。辅助生殖技术主要包括人工授精和体外授精－胚胎移植。

遗传筛查是依据某些指标对群体的某种遗传病进行普查,初步筛选出可疑人群,然后再进一步确诊的过程,包括新生儿筛查和携带者筛查。新生儿筛查有助于早期发现遗传病并行早期治疗;携带者筛查可筛查出群体中的遗传携带者,特别是隐性遗传病杂合子,对预防遗传病在群体中的发生,降低群体的致病基因频率有重要意义。

(刘福民 施旭东 孙贝加)

第十六章　人类基因组计划

人类基因组是指人类细胞中约30亿个碱基对序列所储藏的全部遗传信息的总和，包括细胞核基因组和线粒体基因组。其中，细胞核基因组包含了人类基因组99%的基因，结构基因总数为3万~3.5万个；线粒体基因组是一个结构简单的小基因组，仅有37个基因。

1986年3月，诺贝尔奖获得者Dulbecco在《*Science*》上发表了题为《癌症研究的转折点——人类基因组的全序列分析》的短文，在这篇文章中，Dulbecco回顾了20世纪70年代以来癌症研究的进展，指出：包括癌症在内的多种疾病的发生都与基因有直接、间接的关系，只有对人类基因组进行系统的研究，才能对生命进行系统和科学的解码，进而了解和认识生命的起源、种间和个体间存在差异的原因、疾病的产生机制以及长寿与衰老等现象。

第一节　人类基因组计划的诞生

1984年，美国能源部委托White和Mendelsonhn主持召开一个小型研讨会，讨论测定人类整个基因组的DNA序列的前景和意义。1986年3月，在新墨西哥州的圣菲(Santa Fe)讨论了该计划的可行性，随后美国能源部宣布人类基因组启动计划。1986年，著名的遗传学家McKusick提出从整个基因组的水平上研究遗传的科学称为基因组学。1987年初，美国能源部和国家健康研究院为人类基因组计划(HGP)下拨了约550万美元的启动经费，并开始筹建HGP实验室。1988年2月，美国国家科学研究委员会的专家撰写了《人类基因组的作图与测序》的报道。第2年，美国人类基因组研究中心宣布成立。

1990年10月1日，美国国家卫生研究院和美国能源部联合提出的美国HGP正式启动，其总体计划是在15年内投人至少30亿美元进行人类全基因组序列的分析。随后，欧共体及日本、加拿大、澳大利亚等国也相继提出了各自的HGP，并建立了国际协作的人类基因组组织(HUGO)，以协调国际间的合作。

我国的HGP1994年启动，1999年9月1日，在伦敦举行的人类基因组测序战略第5次会议上，我国被接纳为新成员，成为第6个HGP的成员国，承担人类基因组全部序列1%的测定工作，即3号染色体短臂上的约3 000万个碱基对的测序。做为惟一的发展中国家的中国，加人国际HGP，改变了国际HGP的格局，提高了国际人类基因组计划的国际合作形象，带来了国际社会对国际HGP的支持。2002年5月，联合国教科文组织关于人类基因组基本信息免费共享的声明就是在中国代表的直接努力下促成的。

第二节　人类基因组计划的内容

随着多学科的相互渗人，HGP 成为生命科学领域的第一项大科学工程，它的正式启动标志着解码生命的真正开始，它的最终目标是对整个基因组的 3×10^9 碱基对序列进行测序、作图、基因定位和主要基因的功能分析，为全面认识和了解人类基因组的结构和功能提供详尽的基础资料。HGP 着眼于整个基因组的所有基因，从整体水平上考虑基因的存在、基因的结构与功能以及基因之间相互关系的研究，要达到这些目的，需要进行遗传图、物理图、转录图的绘制和基因鉴定分析等工作。简言之，HGP 的内容可以用 4 张图谱加以概括：遗传图谱、物理图谱、转录图谱和序列图谱。

一、结构基因组学

（一）遗传图谱

遗传图谱（genetic map）又称连锁图谱（linkage map），是以多态性遗传标记为“位标”、通过遗传图距建立起来的、表示基因与遗传标记在染色体上相对位置的基因组图谱，即基因在染色体上相对位置的排列图谱。遗传学的本质就在于揭示生物遗传现象与分子本质之间的关系，遗传图谱就在这个过程中起到桥梁和纽带的作用。它的构建是人类基因组研究必不可少的一步，对于搞清楚基因的功能、定位、分离克隆新基因以及研究染色体上基因的排列顺序起到不可估量的作用。

绘制连锁图谱的基本原理是以染色体上已经定位的基因作为遗传标记，确定与该基因连锁的某性状的基因座位，通过连锁－关联分析，将编码该性状的基因定位于染色体上与遗传标记相近的特定位置上。例如，在 ABO 血型基因中，基因 I^A 位于 9q34，该基因决定血型抗原 A 的存在，进而决定个体血型为 A 血型或 AB 血型。由于 ABO 血型的广泛存在，因此可以用它做为遗传标记，当观察到某一家族中甲－髌综合征（NP）与 A 血型常常伴随出现时，可以认为 I^A 基因和 *NP* 基因表现为连锁遗传，然后，调查某家族后代中只有甲－髌综合征或 A 血型的个体数，这是 I^A 基因和 *NP* 基因之间发生交换的结果。因为 2 个基因之间的距离越远，就越容易发生交换，交换值也就越大，所以如果调查得知只有任一表型的个体数占总个体数的 10%，那么就认为这 2 个基因之间的距离是 10 cM。根据基因之间的连锁－交换重组率来确定 2 个基因之间的连锁距离是传统的遗传依据，但此方法难以做到将基因精细定位，更谈不上如何克隆这一基因。

仅用已知定位的少数几个基因作为遗传标记是难以绘制完整的遗传连锁图谱的，人类基因组 DNA 中存在大量的“微卫星”，它是长 2～6 个碱基对的序列，在染色体某一区域内可以重复几次到几十次，所以又称为短串联重复（STR）。短串联重复在染色体上呈分散分布，平均 100～200 kb 就有 1 个 STR 位点，而且不同的个体之间 STR 的重复次数有所差别，类似的这种序列差异使得高分辨率的现代遗传连锁图谱的构建成为可能。例如，在人类基因组的第 1 个五年计划（1993～1998 年）预定的遗传图谱的平均分辨率为 2～5 cM，实际上 1996 年绘制完成的遗传连锁图谱的相邻遗传标记间的平均距离仅为 0.7 cM（图 16－1）。

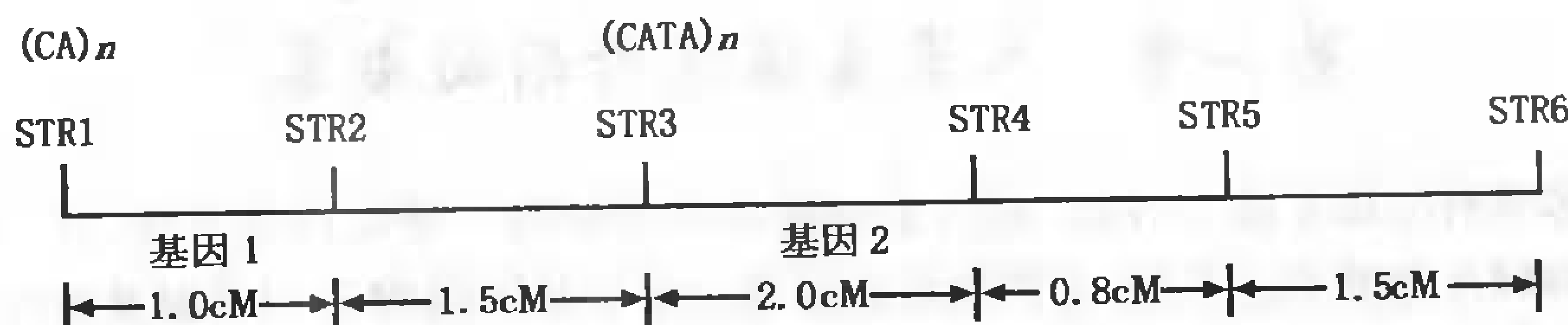


图 16-1 以 STR 为遗传标记绘制的一段连锁图

cM:厘摩,1 cM 相当于 10^6 bp,人类基因组全长约 3 600 cM

现代连锁图谱的概念最早是在 1978 年由 Botstein 提出的,当时是由于 DNA 限制性内切酶和连接酶的应用,限制性片段长度多态性(RFLP)成为一种新的 DNA 多态性标记。RFLP 在基因组分析上有很大的用处,与传统的遗传标记相比较,它具有以下优点:①无表型效应,RFLP 标记的检测不受环境条件和发育阶段的影响;②RFLP 标记在等位基因之间是共显性的,不受杂交方式的影响;③在非等位的 RFLP 标记之间不存在上位效应,因而互不干扰;④RFLP 标记源于基因组 DNA 的自身变异,在数量上不受限制。这些优点使得 RFLP 尤其适合遗传图谱的构建。在 1980 年,Gusella 等在《*Nature*》上报道的人类亨廷顿舞蹈病成为第 1 个被确定为与某个 RFLP 标记连锁的常染色体遗传病,Donis - Keller 等人于 1987 年建成了人的第 1 张 RFLP 图谱。

在人类遗传图谱绘制中使用的遗传标记越多,各个标记的多态性就越高,所得到的遗传连锁图的分辨率就越高。扩增片段长度多态性(AFLP)、随机扩增多态 DNA(RAPD)和单核苷酸多态性标记(SNP)等技术的应用使基因连锁图谱的分辨率越来越高。

(二)物理图谱

基因组中已知的 DNA 序列片段称为序列标记位点(sequence - tagged site,STS),作为 STS 需具备 2 个条件:①在基因组中有明确的位置,②该序列是已知的。这样,以 STS 作为位标,表示基因与 STS 在染色体上的相对位置的基因组图就是物理图谱(physical map)。

物理图谱和遗传图谱的差异表现在作图方法上的不同:遗传图谱是依据家系中基因的连锁、交换绘制的。例如,基因 *AB* 之间是连锁基因,没有交换时基因型为 *AB/ab*,当出现 *Ab/aB* 的基因型个体时,就是连锁基因重组交换的结果,进而就可以根据 *Ab/aB* 基因型个体确定基因之间的交换比值(假设为 8%),我们就用该重组值表示 *A* 基因和 *B* 基因间的遗传图距(8 cM);现代物理图谱是应用多种实验手段,如染色体原位杂交、体细胞杂交、YAC、染色体步行法和 DNA 测序等,确定基因在染色体上的相应位置,用具体的物理单位 Mb 或 kb 表示基因间的距离。根据基因间的连锁交换值进行遗传作图是基因传统的定位手段,这种分析需要在几个家系中完成,而且因为样本采集的随机性导致遗传图谱按重组值来估计实际距离会有较大的偏差,例如在人的 21 号染色体长臂的物理学图距与通过重组率得到的遗传图距之间的误差达 6 倍之多。

早期的物理图谱实际上是细胞遗传学图,仅通过原位杂交将基因定位于染色体区带上,随着研究方法、实验手段的进步,物理图谱越来越精细,现代物理图谱是大量以具有位点专一性的 STS 作图。最常用的寻找重叠克隆的方法是染色体步移法(chromosome walk-

ing method),其原理是筛选单个克隆插入片段作为探针与克隆库其他的克隆进行杂交以发现与它有重叠的克隆插入片段(图 16-2)。

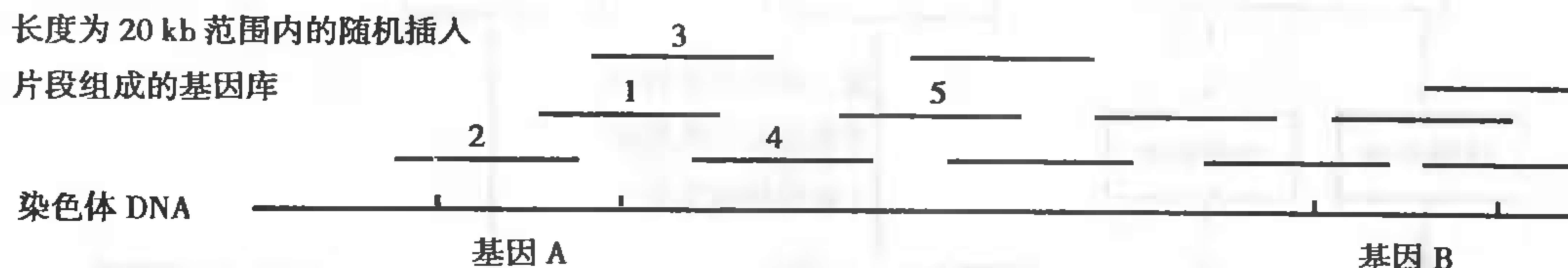


图 16-2 染色体步移法基因定位

在图 16-2 中,如果要从克隆中得到已知遗传距离的基因 *B* 序列,在无特异性探针的情况下,可以应用染色体步移法。这时在附近的基因 *A* 的片段 1 为已知序列,同时,基因 *B* 附近的序列也曾被探针原位杂交鉴定过,并且在大的随机基因组 DNA 库中存在许多重叠片段,因此通过克隆 1 作为探针可鉴定出与之重叠的克隆 2、3 和 4,并通过克隆 4 可以鉴定出克隆 5,依次类推,可沿着这些基因库追查到基因 *B*。

物理图谱的意义在于 STS 能够把经典遗传学与细胞遗传学的位点信息转化为基因组位点的物理信息,基于 STS 位点信息的相邻片段群又提供了研究区域的实验材料,以这些片段的材料便可进行基因组的研究或在该区域寻找新的基因。

(三)转录图谱

生物的性状都由结构或功能蛋白质决定的,而蛋白质又都是由 mRNA 按照遗传密码子编码的,如果能把 mRNA 分离、定位,再转录出 cDNA,这就是基因的主要部分。其中,基因的 cDNA 片段又称表达序列标签(expressed sequence - tag, EST)。由于 cDNA 具有组织、生理以及发育的特异性,所以 EST 除提供序列信息之外,也提供了该基因表达的组织、生理状况、与发育阶段的信息。绘制一张反映正常、异常或受控条件下所有基因的时空图,就可以了解某一基因在不同时期、不同组织中不同水平的表达,也能够了解到某一组织中不同时期、不同基因、不同水平的表达,还可以了解到某一特定时期不同组织中奢侈基因的不同表达水平。人类的疾病(包括对疾病的易感性)涉及基因的结构和表达的改变,有了一张正常的转录图,就奠定了进一步绘制特定的生理、病理、受控条件下的转录图。

从表型到多肽再克隆基因的传统的疾病基因定位、鉴定和克隆方法发展为"位置克隆法"(图 16-3)。

HGP 的开展在基因的定位、测序和功能研究等方面积累了大量的数据,充分利用这些已有的数据资料,可加速基因的克隆。下面介绍序列的同源比较在新基因克隆方面的应用。首先,以已知基因 cDNA 序列对 EST 数据库进行 BLAST 分析,找出与已知基因 cDNA 高度同源的 EST;其次,用相应的软件构建重叠群,并找出重叠群的一致序列;然后,比较各重叠群的一致序列与已知基因的关系,通过 PCR 在 cDNA 文库中将重叠群之间的区域扩增并测序;再次,对编码区蛋白质序列进行比较,与已知基因蛋白质的功能域比较分析新基因的功能;最后,用新基因的序列对 STS 数据库进行 BLAST 分析,如果某一 EST(非重复序列)与某一 STS 有重叠,那么 STS 的定位即确定了新基因的定位。

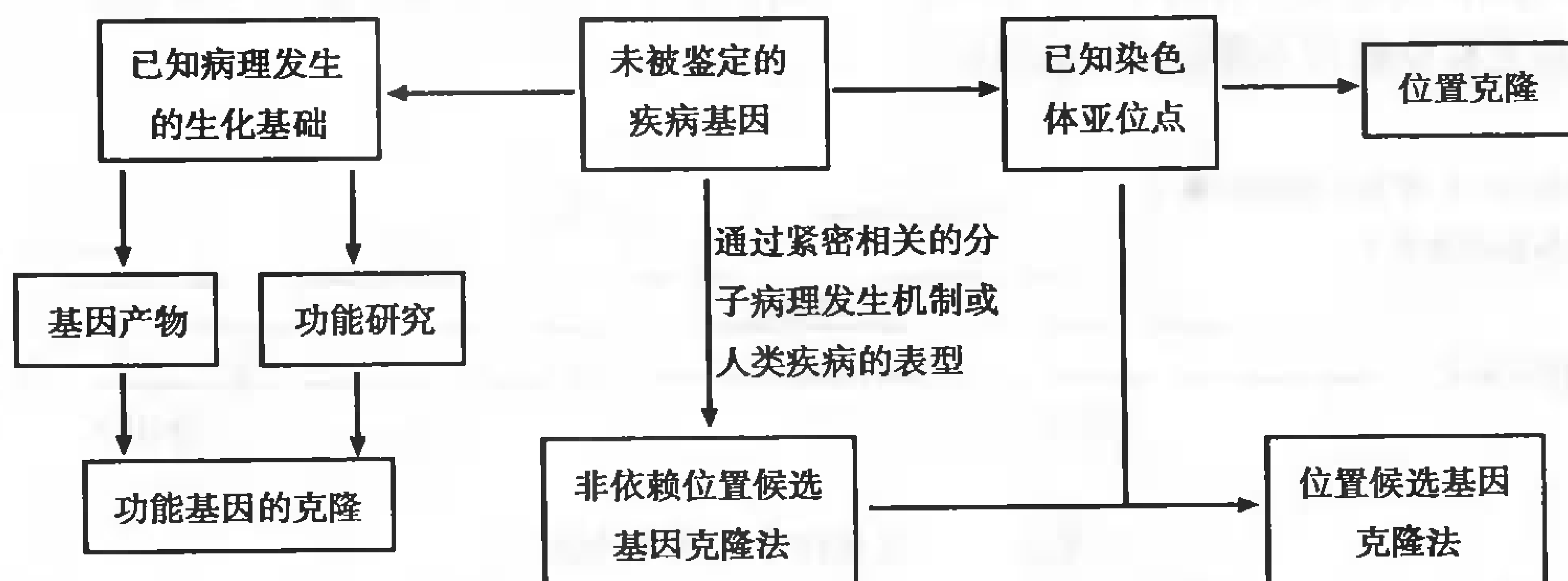


图 16-3　寻找疾病基因的经典途径和主导途径

(四)序列图谱

HGP 的重中之重就是要获得人类基因组序列图，而序列图谱实际上就是要阐明组成人类染色体的 DNA 分子中 4 种脱氧核苷酸(A,T,G,C)的排列顺序。因此，人类基因组核苷酸序列图是分子水平上最高层次的、最详尽的物理图。

按 1995 年预定的 HGP 的目标，2001 年年底前完成人类 1/3 序列的测定，2005 年年底前完成全序列的测定工作，实际上在 2000 年 6 月就基本完成了“工作框架图”。“工作框架图”的构建包括了两大任务：首先，是要确定几十万个大片段 DNA 克隆在人类染色体上的位置；然后，测定这些大片段的化学组成单位——4 种核苷酸的序列组成，完成序列图谱的构建。组装好的“工作框架图”已经能够覆盖人类基因的 97%，其中至少 85% 的序列已经组装得准确无误，20% 的序列已经达到“完成图”的标准，总体质量远超过原来的估计。这项工作之所以能够提前完成，很大程度上取决于 DNA 测序技术的快速发展，随着工艺水平的不断提高、新技术的不断出现，使得 DNA 测序技术日趋规模化和工业化。DNA 测序的规模化的基本趋势之一是技术的并进，除了测序技术之外，PCR 技术、序列分析、DNA 制备、数据的分析和组建等都得到了平行的发展，这些方面的快速发展也使得测序工作的自动化紧随其后，同时，也大大降低了测序的成本。从这个意义上讲，HGP 的序列图谱正是生物技术和计算机技术紧密结合并发展到一定阶段所带来的必然结果。

HGP 提供的遗传图谱、物理图谱、转录图谱和序列图谱，组成了不同层次的、并最终从分子水平上阐明人类遗传信息的“第二解剖图”。当前，人类基因组全序列图实际上是一个“有代表性的人类个体”的序列图，因为所有个体的基因位点都是相同的，不同种族、不同个体的基因差异，以及“正常基因”与“致病基因”的差异，只是等位基因之间在某一位点上的差异。2001 年 2 月，公布了人类基因组测序的初步报道，结果显示人与人之间共享 99.9% 的遗传信息，个体之间的差异仅为 0.01%；编码蛋白质的 DNA 序列仅占基因组的 1.1% ~ 1.4%，其余的为非编码序列。所以，该“有代表性的人类个体”的基因与序列图在理论上可代表全人类的基因组信息，可以用于任何民族群体、任何个体的基因分析及基因诊断。

二、功能基因组学

HGP 的第一阶段主要进行结构基因组学(structural genomics)研究,然而,仅破译人类基因的核苷酸序列对了解生命的本质还远远不够,HGP 的最终目的是要了解人类 3 万～3.5 万个结构基因的功能及表达调控机制,进一步揭示人类生命活动规律。因此,一个以研究基因功能为主要内容的后基因组时代即功能基因组学(functional genomics)时代已经到来。

功能基因组学研究是通过分析基因表达的蛋白质或观察基因表达被阻断后在细胞和整个水平所产生的表型变化来识别基因的功能,这项工作是在完成基因组结构分析后的必然延续,同时也是基因组结构分析的最终目的。这期间主要阐明由基因组内基因编码产生的、真正执行生命活动的全部蛋白质的表达规律和生物意义,以基因产物的结构和功能研究以及开发应用为主要研究方向,揭示基因内核苷酸序列所蕴藏的生物学功能和意义。在后基因组时代,基因功能研究的主要技术和方法包括:转基因技术、反义核酸技术、基因敲除、基因嵌入技术、蛋白质组研究、蛋白质芯片技术和生物信息技术等,这些技术可利用结构基因组学提供的信息,以高质量、大规模的实验方法以及计算机分析为特征,全面、系统地分析基因的功能,这些功能包括生物学功能、细胞学功能及发育学功能等。

基因的功能研究是人类后基因组研究的关键,它以揭示基因组的功能及调控机制为目标,其核心科学问题主要包括:基因组的多样性、遗传疾病产生的起因、模式生物基因组研究、基因的表达调控的协调作用以及蛋白质产物的功能等多个领域。功能基因组学的研究将使人们更深入地理解人类基因组遗传语言的逻辑构架、基因结构与功能的关系。长期以来,个体发育、生长、衰老和死亡机制,神经活动和脑功能表现机制,细胞增殖、分化和凋亡的机制,信息的传递和作用机制,疾病发生、发展的基因及基因后(发病机制、病理过程)机制等诸多方面的问题一直困扰着人们,而功能基因组学的研究给这些问题的进一步解决提供理论基础。在医疗卫生方面,研究成果可用于医药的研究和开发;致病基因或疾病易感基因的鉴定和克隆,可用于遗传病的诊断、治疗和预防方法的设计;医生将能够根据患者的个人遗传构成,进行更加个人化的药物疗法;科学家们在人体器官和组织“重造”以及修复方面将取得巨大进步。以基因组研究成果为基础的基因组工业,将带动一批高新技术产业向新的领域开拓,在农业、畜牧业方面,可用新的方式对动植物疾病进行诊断和处治,改善改良农作物的品质,提高产量;在纺织业及废物处理和环境保护方面,基因组的研究成果也都将发挥极其重要的作用。

功能基因组学延伸的内容有:人类基因组多样性计划、环境基因组学、肿瘤基因组解剖学计划及药物基因组学等。此外,HGP 其延伸内容决定性的成功取决于生物信息学和计算机生物学的发展和应用,主要体现在数据库对数据的储存能力和分析工具的开发,这些都将成为 HGP 延伸篇中的主要内容。

三、环境基因组学

基因的表型效应受环境因素的影响。基因多态性或突变的产生又与环境作用相关,基因与环境的相互作用将对人类产生有害效应。如果遗传易感性已知,就能准确地对引

起疾病的环境成分进行鉴定和评价暴露的准确危险度，也就可以帮助个体准确地认识所处的环境暴露可导致的危险度，进而达到增强对疾病的预防和改善公关卫生的目的。因此，环境基因组计划着眼于研究环境暴露与疾病相互影响。群体流行病学的研究则是鉴定引起疾病的等位片段和环境暴露的相互关系，因此形成了环境基因组计划的主要部分。人类疾病是由先天遗传易感性以及环境因素相互作用导致的，其中，在长期的生活经历中不断积累的环境暴露起到了关键的作用，这一过程非常类似于人类性格的形成过程。个体健康或疾病状态是受多种因素相互作用的结果，如对环境暴露的易感性是其中的重要因素。衰老也同样影响着个体的疾病状态，个体生存越长，则患病概率越大。随着更好地了解环境暴露对人类健康和疾病状态影响的机制，我们就会增加预防测量和改善公共卫生的工作。

一些基因已经被证明在环境暴露的易感性中起重要作用，但这些基因多态性还未被系统地分析、鉴定和报道，环境基因组计划的基本工作就是鉴定这些改变对特定暴露敏感性和抵抗性的基因多样性。环境基因组计划着重研究环境与人类疾病的关系。环境基因组计划将为未来的易感性基因产物和对环境暴露的遗传学反应的分子机制研究提供信息。

四、药物基因组学

药物遗传学是研究包括药物在内的外界化学物质（尤其一些有毒的外源物质）引起机体反应上的遗传多样性。药物基因组学的研究目前正朝向应用生物技术和药物工业化的方法，通过加速鉴定疾病相关靶点以促进药物的发现的方向发展。在基因组上结合化学和高效扫描技术寻找这些靶点，可为发现药物而鉴定更多的新化学物创造条件。以往，90%以上的新化学物进入了临床前实验但未能最终进入临床，导致这样的结果大多是因为在动物实验和部分人群的预实验中显示非显著疗效或者有毒性，而实际上在另外一些人群中所得结果则完全相反。除了具有药物遗传学的研究的遗传多样性引起对药物或有毒物质反应的差异的内容外，药物基因组学还研究包括治疗效果和药物反应不同造成的个体差异，以及每个个体基因组上所存在的与药物作用的不同疾病靶点分子的情况。

影响药物反应多态性的性状结果与RNA和蛋白质的改变有关。转录分析可为药物基因组学提供分析数据，大规模的转录分析的主要结果是产生密集的cDNA和与用荧光或同位素标记的RNA或cDNA的杂交寡核苷酸列阵，通过比较从不同人群中抽提的RNA杂交信号来鉴定基因调节的差异。药物基因组学运用基于列阵的转录分析研究患者组织在对临床性状治疗时的反应，在癌症研究中由于RNA都是来自于活检和手术标本，因此更侧重于表达分析，用以检测肿瘤发生和它们对化学物质治疗反应时的体细胞改变，如原癌基因 $erb-B_2$ 的扩增，可用于预测CMF佐剂对乳腺癌治疗的良性反应。

药物基因组学不同于以往的定位克隆和EST测序等技术，而是开辟了一个崭新的领域，其应用的规模和范围都很大，并且将用于发现临床上具有潜在的效应和毒性资料的化合物。通过药物基因组学可对有限市场的低效和高毒性药物加以改善。随着对人类群体和疾病特定模型及毒性的分子水平的认识，药物基因组学将起到优化药物的作用。

第三节 人类基因组计划的意义

HGP 的实施把人类带入了以生物信息学为主导的生物科学和生物技术的新时代，对于人类认识自身，提高健康水平，推动生命科学、医学、生物技术、制药业、农业等多方面的发展都具有极其重要的意义。其具体作用表现在以下几个方面。

（一）鉴定人类全部基因，揭开生命的奥秘

HGP 之前的遗传学偏重于单个基因的研究，而 HGP 则是着眼于整个基因组的所有基因，从整体水平上考虑基因的存在、基因的结构与功能以及基因之间的相互关系，它将从整体上分析人类的起源、进化、迁徙与分子多态变化事件的相互关系，为生物起源、进化规律的探索提供证据。

（二）把人类带入基因医学的新时代

随着序列图谱的完成，许多致病基因被确定，并逐步对其结构、功能进行了分析，多基因遗传病也已经成为疾病基因组学的研究重点，引发多基因病的数个基因的 SNP 及其特定组合可能是造成疾病易感性的重要原因。因此，个体之间的遗传多态性导致不同个体对药物反应的差异性，这就要求现代医学推进基因治疗方案的个体化。

（三）带动生物制药产业迅速崛起

HGP 的实施使人类在了解致病遗传机制和发现新基因上迈出了重要的一步，这为基因药物的设计提供了理论基础和设计原则。由于基因药物的专一性和高效性将给商家带来巨大的产业利润，因而，基因制药产业将成为未来经济的重要支柱产业之一。

（四）对相关生物技术产业产生有力的推动作用

现代生物技术涉及到医疗卫生、制药、环境改造、食品改良等诸多领域，HGP 的大规模运作也将推动生物技术科研与开发并肩走向规模化和自动化。因此，这无疑会使生物技术在未来的经济发展中占有及其重要的位置。

（五）促进学科间的交叉和重组

HGP 的深入发展诞生了许多新学科、新领域，包括近几年才提及的生物信息学、比较基因组学、蛋白质组学、医学基因组学和药物基因组学等。这些学科的兴起不是单一学科的简单推进，而是要求多元学科的综合发展，这就为学科间的组合提供了广阔的发展空间。

人们对真理的追求是没有止境的，一个科学问题的解决往往又给另一个科学问题的提出和解决提供了新的思路和方法。HGP 也正是如此。完成人类基因组测序后意味着结构基因组学的结束，随之而来的是功能基因组学时代。有人将 HGP 比作生命周期表，因为它不再是从个别基因的研究着手，而是力求在细胞水平解决基因组问题，同时研究所有基因及其表达产物，以建立对生命现象的整体认识。

当然，HGP 同样也是一把“双刃剑”，它在给人类社会带来巨大发展潜能的同时，也涉及了伦理学方面的问题。基因识别和基因功能鉴定等研究将给人类带来极大的好处，但 HGP 面临着许多严峻的伦理学问题。比如遗传信息的隐私权问题，基因图谱和信息的使

用与人的社会权利问题，基因组信息的医学解释与心理、名誉损害问题，基因资源外流、垄断问题，DNA 银行管理问题等。它的伤害对象可能是个人、家庭以及群体。例如，泄漏了一个人的基因型后可能导致他的就业困难，也可能由于某个隔离种群存在基因缺陷导致他们得不到公平的对待。因此，HGP 不仅深入到人的生命本质，而且涉及到人的价值观，也会改变我们的伦理观念。

小　结

HGP 是 1990 年正式启动的一项多边合作计划，包括 2 个独立又相互联系的研究阶段：结构基因组学和功能基因组学。

结构基因组学是 HGP 的第一步，主要目标是完成 4 个图谱的绘制：遗传图谱、物理图谱、转录图谱和序列图谱。遗传图谱是表示基因与遗传标记在染色体上相对位置的基因组图谱；物理图谱是表示基因与序列标记位点在染色体上的相对位置的基因组图谱；转录图谱是通过表达序列标签来反映基因的表达情况的图谱；序列图谱是组成人类染色体的 DNA 分子中 4 种脱氧核苷酸的排列顺序的图谱。

目前，结构基因组学的研究已基本完成，人类基因总数为 3 万 ~ 3.5 万，仅占基因组的 2%左右。功能基因组学主要通过分析基因表达的蛋白质或观察基因表达被阻断后在细胞和整体水平所产生的表型变化来识别基因的功能。此外，还包括许多延伸内容的研究。

HGP 对于揭示生命的本质，推动生命科学、医学等多方面的发展都具有极其重要的意义。它促进多学科的融合，带动相关产业的发展，推动生物技术及其相关技术的快速发展，使 21 世纪成为生命科学的世纪。

（钱　刚）

英汉名词索引

(按英文字母排序)

参考文献

1 陈诗书,汤雪明.医学细胞与分子生物学(第2版).北京:科学出版社,2004
2 李 璞.医学遗传学.北京:北京大学医学出版社,2003
3 龚非力.医学免疫学.北京:科学出版社,2003
4 王金发.细胞生物学.北京:科学出版社,2003
5 杨抚华.医学生物学(第5版).北京:科学出版社,2003
6 余元勋,等.中国遗传咨询.安徽:安徽科学技术出版社,2003
7 黄荷凤.现代辅助生育技术.北京:人民军医出版社,2003
8 陆国辉.产前遗传病诊断.广州:广东科技出版社,2002
9 邓耀祖,屈伸.医学分子细胞生物学.北京:科学出版社,2002
10 刘建欣,郑昌学.现代免疫学.北京:清华大学出版社,2002
11 温进坤,韩梅.医学分子生物学——理论与研究技术.北京:科学出版社,2002
12 高晓明.医学免疫学基础.北京:北京医科大学出版社,2001
13 T.D.盖莱哈特,等.医学遗传学原理.北京:科学出版社,2001
14 王培林.医学遗传学.北京:科学出版社,2001
15 蔡绍京,徐 珊.医学遗传学.北京:科学出版社,2001
16 左 伋.医学生物学(第5版).北京:人民卫生出版社,2001
17 陈 竺.医学遗传学.北京:人民卫生出版社,2001
18 王培林,傅松滨.医学遗传学.北京:科学出版社,2001
19 傅松滨.医学遗传学.北京:人民卫生出版社,2001年
20 冯作化.医学分子生物学.北京:人民卫生出版社,2001
21 陆振虞.医学遗传学(第2版).上海:上海科学技术文献出版社,2001
22 陈慰峰.医学免疫学.北京:人民卫生出版社,2001
23 孙树汉.基因工程——原理与方法.北京:人民军医出版社,2001
24 卢惠霖,卢光秀.人类生殖与生殖工程.郑州:河南科学技术出版社,2001
25 王培林.遗传医学.北京:人民卫生出版社,2000
26 吴 刚,伦玉兰.中国优生科学.北京:科学文献技术出版社,2000
27 贺 林.解码生命.北京:科学出版社,2000
28 王培林.遗传病学. 北京:人民卫生出版社,2000
29 周宏灏.分子药理学.哈尔滨:黑龙江科学技术出版社,1999
30 魏尔清.药理学前沿——信号、蛋白因子、基因与现代药理.北京:科学出版社,1999
31 李 璞.医学遗传学.北京:北京医科大学、中国协和医科大学联合出版社,1999
32 陆莉娜.高新科技在医学领域的应用(上).北京:长征出版社,1999
33 左 伋,黎立瑾.医学细胞生物学.上海:复旦大学出版社,1998
34 赵 刚.医学遗传学教程.北京:科学出版社,1998

35 柳家英.医学遗传学.北京:北京医科大学出版社,1998
36 吴乃虎.基因工程原理(第2版).北京:科学出版社,1998
37 左 伋,张克雄.医学遗传学.上海:上海医科大学出版社,1998
38 谷志远.现代医学分子生物学.北京:人民军医出版社,1998
39 卢大儒,邱信芳,薛京伦.医学分子遗传学.上海:复旦大学出版社,1998
40 王亚馥,戴灼华.遗传学.北京:高等教育出版社,1998
41 杨抚华.医学生物学(第4版).成都:四川科学技术出版社,1997
42 彭奕欣,黄诗笺.进化生物学. 武汉:武汉大学出版社,1997
43 陈仁彪,冯波.医学遗传学.上海:上海科学技术文献出版社,1997
44 左 伋.医学遗传学.上海:上海医科大学出版社,1997
45 杜传书.医学遗传学基础(第2版).北京:人民卫生出版社,1996
46 刘权章.临床遗传学彩色图谱.北京:人民卫生出版社,1995.
47 细胞生物学名词审定委员会.细胞生物学名词.北京:科学出版社,1992
48 杜传书、刘祖洞.医学遗传学(第2版).北京:人民卫生出版社,1992
49 李宝森,胡庆宝.遗传学.天津:南开大学出版社,1991
50 生物化学名词审定委员会.生物化学名词.北京:科学出版社,1990
51 遗传学名词审定委员会.遗传学名词.北京:科学出版社,1989
52 夏家辉,等.染色体病.北京:科学出版社,1989

书名
前言
目录